

令和 6 年 6 月 11 日現在

機関番号：21201

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12143

研究課題名（和文）時・場所・状態を考慮した社会課題解決につながるWebパーソナライズ技術の提案

研究課題名（英文）A Proposal for Web Personalization Technology that Leads to Social Problem Solving by Considering Time, Place, and Condition

研究代表者

堀川 三好（Horikawa, Mitsuyoshi）

岩手県立大学・ソフトウェア情報学部・教授

研究者番号：40337473

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究は、モバイル端末利用者を対象に利用者の状態を考慮するWebパーソナライズの実現を目指している。まず、Webブラウザから取得される加速度データに変化点検知を適用することで、利用者の状態区間推定を行う手法を提案し、検証実験により有効性の確認を行った。次に、各状態区間における動作状態と利用状態から状態推定を行う手法を開発した。提案手法は、学習段階として対象Webサイトにおける状態をラベル付けする状態定義と、推定段階としてセンサデータを入力とした機械学習による状態分類で構成される。その結果、即時性の高い状態分類が可能であり、Webパーソナライズへの活用が可能であることを明らかにした。

研究成果の学術的意義や社会的意義

最近のWebパーソナライズをめぐる動向として、サードパーティCookieの廃止・規制がある。従来は、この技術を用いることで、異なるWebサイト間で利用者の情報や行動を追跡可能とし、特に、ターゲティング広告の分野でユーザの関心に基づいた広告配信を実現していた。しかしながら、プライバシーを侵害しうることに対する懸念が広がり、Cookieを廃止・規制する動きが加速している。今後は、プライバシーを保護しつつサードパーティCookieに頼らない新たなWebパーソナライズを実現することが求められている。本研究の成果は、新たなWebパーソナライズを実現するための独自性高い技術として学術的・社会的意義がある。

研究成果の概要（英文）：This research aims to realize web personalization that considers the user's state, targeting mobile device users. First, we propose and verify the effectiveness of a method for estimating user state intervals by applying change point detection to acceleration data obtained from web browsers. Next, we developed a method for estimating the state from the motion state and usage state in each state interval.

The proposed method consists of a state definition that labels the state of the target website in the learning phase, and a state classification by machine learning using sensor data as input in the estimation phase. As a result, it was clarified that state classification with high immediacy is possible, and it can be applied to web personalization.

研究分野：社会システム工学

キーワード：Webパーソナライズ センサデータ モバイル端末 状態推定

## 1. 研究開始当初の背景

モバイル端末の爆発的な普及により、利用者の情報を収集して一人ひとりに合わせてコンテンツの最適化を図るパーソナライズの導入が身近なものとなった。しかしながら、さらなるパーソナライズの進化には、時・場所に加え状態を把握することが重要になると考えている。併せて、モバイル端末の高性能化に伴い Web ブラウザを用いて場所情報、センサ情報や操作情報が取得可能となった。特に、Web で利用可能な JavaScript による機械学習ライブラリが提供され始め、汎用的な利用が可能な Web アプリにおける状態推定を用いたサービスの広まりが期待できる。

## 2. 研究の目的

本研究は、モバイル端末利用者を対象に時間、位置情報、検索・閲覧・購買履歴および利用者属性に加えて、利用者の状態を考慮する Web パersonナライズの実現を目指している。ここで、本研究における利用者の状態を「Web 閲覧時の利用者の現実空間での動作状態とアクセス解析から得られるサイバー空間の利用状態を考慮したもの」と定義する。

## 3. 研究の方法

利用者の状態を把握するためには、単なる動作推定のみではなく、一連のものとみなすべき状態が連続した区間（以降、状態区間）を検出する必要がある。まず、Web ブラウザから取得される加速度データに変化点検知を適用することで、利用者の状態区間推定を行う手法を提案し、検証実験により有効性の確認を行う。

次に、各状態区間における動作状態と利用状態から状態推定を行う手法を開発する。提案手法は、学習段階として対象 Web サイトにおける状態をラベル付けする状態定義と、推定段階としてセンサデータを入力とした機械学習による状態分類で構成される。

## 4. 研究成果

### (1) 状態区間推定

Web ブラウザから取得される加速度データに対して変化点検知を適用することで、利用者の状態区間推定を行う。本研究では、変化点検知手法の 1 つである Change Finder を利用する。変化点検知の閾値の設定にはボリンジャーバンドを利用する。また、特性変化点という概念を組み合わせることで、Web ブラウザ上で行う計算の回数を抑制する手法を提案する。以下に述べる手法をプライバシー保護の観点から、モバイル端末の Web ブラウザで計算処理を行うものとする。

#### (1-1) 特性変化点

Web ブラウザで変化点検知を行う場合、得られたデータを逐次処理すると計算回数が膨大になる。そのため、状態変化が起こる可能性がある時点（以降、特性変化点）の検出を常時行い、特性変化点が検出された際に変化点検知処理を行うこととする。特性変化点の検出は、位相幾何学のモース理論を用いる。研究[2]では、特性変化点について計算処理量や状態区間算出への影響を検証した。その結果、特性変化点処理を入れることで計算量を約半分に抑制し、状態区間算出への影響が小さいことを示した。

#### (1-2) 変化点スコア

本研究では、変化点検知手法として Change Finder を用いる。Change Finder は 2 段階にわたって時系列モデルを学習させることで、データの時系列的変動を示す変化点スコアを算出する手法である。本研究では、変化点検知された時点から次の変化点検知までを状態区間として扱う。

#### (1-3) ボリンジャーバンド

ボリンジャーバンドは、時系列データの移動平均値におけるばらつきを標準偏差で表すことで、異常値を捉える手法である。平滑化区間の移動平均値および標準偏差から変化点スコアの正常状態の範囲を算出し、上限または下限を超えた場合に変化点として検出する。

#### (1-4) 3.1.4 変化点検知のパラメータ決定方法

上述の Change Finder とボリンジャーバンドを用いる場合、各パラメータ値を適切に設定する必要がある。Change Finder では、パラメータとして AR モデルの次数、忘却率および平滑化区間の長さを決定する必要がある。また、ボリンジャーバンドのパラメータとして、標準偏差を設定する必要がある。研究[3]では、複数の実験条件下で行動した際のセンサデータを収集し、パラメータの設定方法を検討した。その結果、モバイル端末保持動作における変化点検知においては、AR モデルの次数、忘却率や平滑化区間の長さは、個人特性によらず固定値で設定することができ、ボリンジャーバンドの標準偏差については、個人特性に応じて設定する方針を得た。

### (2) 状態推定手法

提案する状態推定手法は、状態定義および状態分類の 2 つで構成される（図 1）。上述の状態区間推定で検出された区間に対して、状態定義では、学習段階として動作状態や利用状態からクラスタリングを行い状態のラベル付けを行う。動作状態や利用状態を考慮した場合、どのような状態があるかは、対象とする Web サイトのコンテンツに依存すると考えられる。そのため、状態区間ごとに動作状態と利用状態を考慮したクラスタリングを行うことで、どのような状態が生じているかを定義する。状態分類では、推定段階として即時性高い入力データを用いて、状態定義で算出されたどの状態に分類されるかを算出する。入力データとして加速度・角速度のセンサデ

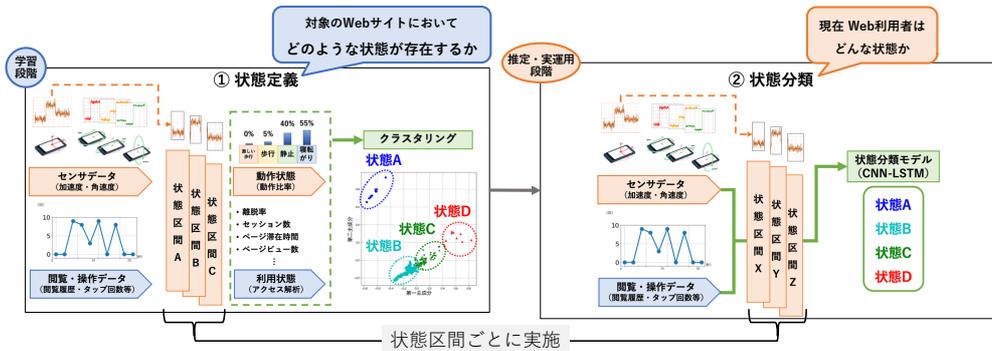


図1 状態推定手法の概要

ータおよび閲覧履歴等のデータを用いる。

(2-1) 状態定義

Web 利用者の現実空間での動作状態を表現するために、本研究では状態区間ごとに利用者の動作比率を算出する。動作比率とは、状態区間内で行われた歩行・激しい歩行・静止・寝転がりの4つの動作の割合を意味する。この算出のため、研究[1]で提案した動作推定モデル (CNN-LSTM) を用いる。動作推定モデルは、F 値 0.90 を示しており、JavaScript と TensorFlow.js により Web ブラウザで処理可能なプロトタイプを開発した。

(2-2) 利用状態

Web 利用者のサイバー空間での利用状態を表現するために、ページビュー数や離脱率等のアクセス解析データを算出する。本研究では、セッションまたは状態区間ごとに算出可能なアクセス解析データを取り上げる。セッションごとのアクセス解析は、一般的なアクセス解析で利用される基本的なデータを取り上げる。状態区間ごとのアクセス解析は、従来の Web にはないセッションを更に現実空間の動作で細分化したものとなる。

(3) 検証実験

(3-1) 実験概要

状態定義において動作比率とアクセス解析データを用いたクラスタリングにより、状態ラベルの定義が可能かを明らかにする。特に、セッションごとの分析と状態区間ごとの分析の比較を行う。加速度・角速度等のセンサデータ、アクセス解析データ、状態区間を収集するために HTML と JavaScript で作成した Web アプリを用いる。Web アプリは、複数のニュース記事サイトの URL が設置されているニュースポータルサイトを作成した。センサデータを 200 ミリ秒ごとに収集し、センサデータの収集と同時に CNN-LSTM モデルによる動作推定を行う機能を持つ。アクセス解析データについては、セッションおよび状態区間ごとに算出する。

(3-2) 実験結果

セッションデータと区間データを用いてクラスタリングを実施するために、k-means 法におけるクラスタ数をエルボー法によってクラスタ数を 5 とした。主成分負荷量の結果から、セッションデータの場合はセッション時間やニュース記事サイト・ニュースポータルサイトの閲覧時間など、時間に関する変数と主成分との相関が強く現れている。区間データの場合は、閲覧時間などの変数に関する相関は強く現れてはならず、特定の動作の比率の高さと離脱率の高さに関して相関が現れていた。

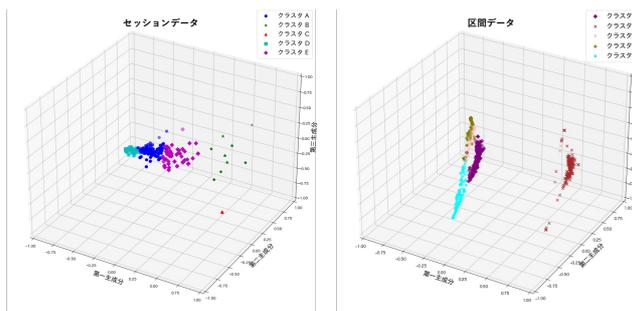


図2 状態定義のためのクラスタリング例

また、定義された5つのクラスタリング結果に対し、センサデータから状態分類する学習モデルを開発した。3分類モデルに対して1秒程度の入力値でF 値 0.90 を実現している。今後は、状態分類モデルについて開発を進めて、状態推定手法全体を社会実装することにより有効性の検証を進める予定である。

<引用文献>

[1]小野寺斗弥, 堀川三好, 猪股一步希. “利用者の時・場所・状態を考慮した Web パーソナライズの提案,” 日本経営工学会 2021 年度春季大会, 2021  
 [2]大澤嘉規, 堀川三好, 岡本東. “モバイル端末を用いた Web パーソナライズのための状態区間検出方法の提案”, FIT2022 第 21 回情報科学技術フォーラム, 2022  
 [3]大澤嘉規, 堀川三好, 岡本東. “変化点検地を用いたモバイル端末利用者の状態推定手法”, 情報処理学会 第 85 回全国大会, 2023  
 [4]大澤嘉規, 堀川三好, 岡本東. “モバイル端末利用者の Web パーソナライズに向けたセンサによる状態推定手法の提案”, 情報処理学会第 109 回 MBL 研究会, 2023

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計6件（うち招待講演 0件 / うち国際学会 0件）

|   |
|---|
| 1. 発表者名<br>大澤嘉規, 堀川三好, 岡本東                      |
| 2. 発表標題<br>モバイル端末を利用したWebパーソナライズのための状態区間検出方法の提案 |
| 3. 学会等名<br>第21回情報科学技術フォーラム講演論文集                 |
| 4. 発表年<br>2022年                                 |

|  |
|--|
| 1. 発表者名<br>大澤嘉規, 小野寺斗弥, 五十嵐大地, 堀川三好, 岡本東 |
| 2. 発表標題<br>変化点検知を用いたモバイル端末利用者の状態推定手法     |
| 3. 学会等名<br>情報処理学会第85回全国大会                |
| 4. 発表年<br>2023年                          |

|   |
|---|
| 1. 発表者名<br>小野寺 斗弥, 堀川 三好, 猪股 一步希                |
| 2. 発表標題<br>利用者の時・場所・状態を考慮したWebパーソナライズの提案        |
| 3. 学会等名<br>日本経営工学会2021年春季大会（Best Paper Award受賞） |
| 4. 発表年<br>2021年                                 |

|   |
|---|
| 1. 発表者名<br>大澤嘉規, 猪股一步希, 堀川三好, 岡本東             |
| 2. 発表標題<br>モバイル端末利用者のWebブラウザを用いた動作・状態推定モデルの提案 |
| 3. 学会等名<br>情報処理学会第84回全国大会（学生奨励賞受賞）            |
| 4. 発表年<br>2022年                               |

|                                    |
|------------------------------------|
| 1. 発表者名<br>猪股一步希, 堀川三好, 岡本東        |
| 2. 発表標題<br>CTGANによるセンシングデータ拡張手法の提案 |
| 3. 学会等名<br>情報処理学会第84回全国大会          |
| 4. 発表年<br>2022年                    |

|   |
|---|
| 1. 発表者名<br>小野寺斗弥, 猪股一步希, 大澤嘉規, 長久保伊吹, 堀川三好, 岡本東 |
| 2. 発表標題<br>動作推定を用いた新たなWebサービスの提案と評価             |
| 3. 学会等名<br>情報処理学会第84回全国大会 (学生奨励賞受賞)             |
| 4. 発表年<br>2022年                                 |

〔図書〕 計0件

〔出願〕 計1件

|                                   |              |               |
|-----------------------------------|--------------|---------------|
| 産業財産権の名称<br>情報処理システム、情報処理方法、プログラム | 発明者<br>堀川三好他 | 権利者<br>同左     |
| 産業財産権の種類、番号<br>特許、2022-055292     | 出願年<br>2022年 | 国内・外国の別<br>国内 |

〔取得〕 計0件

〔その他〕

6. 研究組織

| 氏名<br>(ローマ字氏名)<br>(研究者番号) | 所属研究機関・部局・職<br>(機関番号) | 備考 |
|---------------------------|-----------------------|----|
|---------------------------|-----------------------|----|

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
|---------|---------|