

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 Proposal of a State Estimation Method for Web Personalization Using Mobile Devices

大澤 嘉規[†] 堀川 三好[†] 岡本 東[†]
Yoshiki Osawa Mitsuyoshi Horikawa Azuma Okamoto

1. はじめに

製品やサービスを利用者に併せて提供することを意味するパーソナライズは、マーケティングや情報システム分野において欠かせない技術となっている [1]。特に、Web パーソナライズは、購入・閲覧履歴、利用者属性、時間および場所等の情報から、利用者の意識的・潜在的に必要としている情報を、的確なタイミングで届けることを意味し、Web マーケティングの分野で広く活用されている。例えば、EC サイトのレコメンデーション、ターゲティング広告、利用者の時と場所に応じたニュース提供や検索結果の表示等、多岐に渡り利用されている。また、コンビニの賞味期限切れ間近の見切り品情報を、位置情報と購買履歴からモバイル端末に送信する [2]等の新しいサービスが始まっている。この事例は、パーソナライズ技術が単に顧客サービス向上のみでなく食品ロスの軽減にも貢献し、社会課題解決にもつながることを意味している。

また、最近の Web パーソナライズに関する社会的な動きとして、個人情報に関するプライバシー保護の強化が挙げられる。従来、ターゲティング広告の中心技術になっていたサードパーティ Cookie が利用できなくなりつつあり、プライバシーを保護しながら Web パーソナライズを行う手法が求められている。

本研究は、時・場所や検索・閲覧・購買履歴の情報に加えて、利用者の状態を考慮した Web パーソナライズを実現することを目的としている。ここでの「状態」とは、モバイル端末を保持している利用者が、どのような一連の動作のもとで閲覧しているかを意味している。例えば、同じ時・場所や閲覧履歴を有する利用者でも、駅において電車を待つ状態と移動する状態とでは、Web 閲覧時に配信すべき情報は異なる。移動する状態の場合は、Web ページ内で移動時によく利用される必要最低限の情報のみを提示、または静止するよう注意喚起することで、社会課題の 1 つである駅構内の歩きスマホを最小限に抑えることが可能となる。

本稿では、モバイル端末で Web 閲覧を行う利用者の状態推定を行う手法として、Web ブラウザから取得される加速度、角速度のセンサデータと機械学習を用いる。まず、取得されるセンサデータから状態区間推定を行う。人間の行動における「改札を通過する」、「ホームへ向かっている」「電車を待っている」等の一連の動作のまとまりである状態を、連続時間を区切った区間として捉えることができる。本稿では、特定の状態が始まってから次の状態に切り替わるまでの時間を、状態区間と定義する。この状態区間を推定するために、加速度データに対して変化点検出を適用することで状態区間推定を行う。また、プライバシー保護を念頭においた Web パーソナライズを実現するため、モバイル端末内で状態推定を即時性高く行い、その結果のみを Web サーバーへ送信して利用する仕組みを目指している。そのため、変化点検出をセンシングの都度行った場合、処

理量が増加することが考えられる。これを回避するために、本研究では、特性変化点を定義し、変化点検出の計算タイミングとして活用する。併せて、状態区間推定を行いつつ、都度収集される加速度・角速度データを用いて動作推定を行う。モバイル端末保持者の動作推定は、既に多くの研究で取り組まれているが、Web ブラウザから取得されたデータにより機械学習を用いて行った事例は少ない。本稿では、小野寺 [3]らの先行研究で開発したモバイル端末操作時の基本動作である「歩く」「静止」「階段」「寝転がる」の 4 つの動作を判別する動作推定モデルを活用する。本稿では、上述の状態区間推定と動作推定を用いることで、利用者の状態を把握することが可能かを検討するために、人間が付けた状態ラベルと提案手法によって導出された状態推定結果の間に、どのような差異が生じるかを、複数のシナリオや利用者特性を比較することで明らかにする。

提案手法による状態推定は、状態に応じたコンテンツ配信、適切なタイミングでの広告配信および状態を考慮したレコメンデーション機能等の幅広い分野での適用が考えられる。これにより、プライバシーを保護しつつ利用者の状態を考慮することで、新たな Web パーソナライズ技術を確立することを目的としている。

2. 関連・先行研究

2.1 Web パーソナライズに関する技術・研究

Salonen ら [1]は、Web パーソナライズに関する研究を幅広く調査し、人間中心アプローチによる研究、レコメンデーションシステム等の実装に関する研究、理論的基礎に関する研究の 3 つに分類している。また、Web パーソナライズと Web カスタマイズの違いについて述べており、パーソナライズがコンテンツを提供する側が顧客にあわせて最適化するのに対し、カスタマイズは利用者自身が設定するものだと述べている。

Web パーソナライズの中でも、レコメンデーションに関する研究は多く行われている。Wu [4]らは、レコメンデーションに関する研究を協調フィルタリング型、利用者属性やアイテムベースの知識グラフを用いたコンテンツ強化型、時間・場所・閲覧履歴にもとづく時系列型の 3 つに分類している。また、Ko [5]らは、技術概要と共に利用されている分野を調査し、EC サイトの他にも SNS、観光、ヘルスケアの分野でよく研究されていることと述べている。

従来、Web パーソナライズ、特にターゲティング広告には、サードパーティ Cookie が広く用いられていた。サードパーティ Cookie とは、訪問している Web サイトと異なるドメインから発行される Cookie を意味し、利用者の Web 上の行動に基づいて、属性や興味・関心度の高さに適した広告の配信を可能とする技術である。しかしながら、プライ

[†] 岩手県立大学大学院ソフトウェア情報学研究科
Graduate School of Software and Information Science, Iwate
Prefectural University

プライバシー保護が強化される中で、サードパーティ Cookie を規制・廃止する動きが加速している。こうした中、Google が開発する FLoC や Topics は、プライバシーを保護しつつ、利用者の興味に基づいた広告配信を行う技術として注目されている [6] [7]。特に Topics は、似たような閲覧履歴を持つ利用者をブラウザ内でグループ化し、グループの ID を Web サーバーに提供することで、ターゲティングを図る仕組みとなっている。グループ化のための処理は、利用者端末内で行われ、個々の閲覧履歴の情報が外部企業に共有されることはなく、プライバシー保護と効果的な Web パーソナライズ手法として開発が進められている。

2.2 状態推定に関する研究

2.2.1 人間の状態分析

Okoshi ら [8] は、動作情報やモバイル端末の操作情報を用いて、人間の状態分析をする手法を提案している。この研究は、利用者の生活の区切り点を特定することにより、モバイル端末などの通知タイミングの最適化を図ることを目指している。これにより、通知によって起こる利用者の精神的負担を抑えることを目指している。生活の区切り点の特定には、モバイル端末上のアプリの使用状況の情報に加え、身体活動の変化の情報を用いている。Wampfler ら [9] は、モバイル端末の操作情報とセンサデータを用いて、人間の感情分析を行う手法を提案している。この研究は、覚醒度合い、ストレス度合いおよび感情などを被験者に自己申告してもらい、モバイル端末のキーストローク情報と加速度、角速度の値との関連性を可視化している。これらの研究によって、動作情報を用いた人間の状態変化の検知手法の有用性が示されている。

大竹ら [10] は、行動の過程や心理的影響を考慮した状態分析手法を提案している。この研究では、加速度の値を用いた動作推定機能に加えて、行動時の心理状況を記録する機能を有するアプリを使用して、利用者の心理状態に合わせた情報提供手法を提案している。これにより、利用者の嗜好や現在の状況を考慮したサービス提供を可能としている。

これらの研究では、人間の状態を単一の動作ラベルで表現している。例えば、「座ってコーヒーを飲んでいる」という身体活動や「歩き疲れている」という心理状況、「10 分間歩き続けている」という行動の過程などが、例として挙げられる。しかし、実際の生活において、同じ動作を行っていても、違う状態である場合が想定される。例えば、「駅のホームを歩いている」状態と「店舗内を歩いている」状態では、同じ「歩く」動作を行っているものの、心理状況や行動の過程は大きく異なる。このような点から、状態の定義には単一の動作ラベルではなく、一連の動作ラベルを考慮することが求められる。

2.2.2 時系列データの状態分析

Atluri ら [11] は、時空間データに対して、有益な情報を取得するための手法を幅広く調査している。その中でも、時系列データの状態変化を検出する手法として、変化点検知が有効であることを述べている。これは、一連の観測データの統計的性質が急激に変化した時点を検出する手法である。例えば、Miralles ら [12] は、大陸における水の平均蒸発量に対して変化点検知を適用し、エルニーニョ現象からラニーニャ現象へ遷移する時点を検出している。これによ

り、大陸における水の循環が加速していることを示し、地球温暖化の抑制の一助となる手法を提案した。この研究事例の他にも、様々な分野において変化点検知の活用が進んでいる。

3. 状態推定手法

提案手法の概要を図 1 に示す。本提案手法は、モバイル端末の Web ブラウザから取得される加速度・角速度データを用いて、利用者の動作を推定する「動作推定」、加速度データを用いて状態変化の検知を行い区間決定する「状態区間推定」、推定された状態区間内における動作推定結果やタップ操作情報を入力値とした機械学習によって、利用者の状態を分類する「状態推定」によって構成される。本稿では、主に「状態区間推定」について報告をする。

3.1 動作推定手法

ブラウザで動作する HTML と JavaScript を用いて加速度・角速度等のセンサデータ、位置情報、タップ等の操作情報を取得することが可能である。小野寺ら [3] は、この技術を用いて取得されるセンサデータから、モバイル端末の操作時の基本動作となる「歩く」「静止」「階段」「寝転がる」の 4 分類を行うために、図 2 のような構造の CNN-LSTM モデルを構築した。このモデルは 0.90 の F 値を示し、図 3 のような混合行列が求められる。

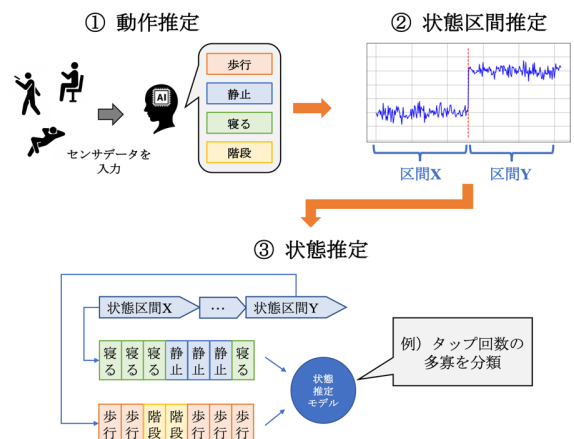


図 1 提案手法の概要

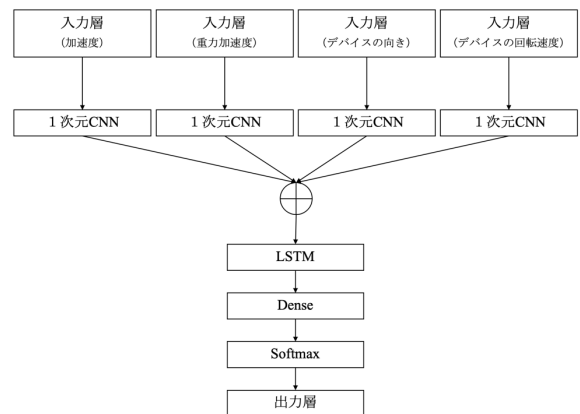


図 2 動作推定モデルのアーキテクチャ

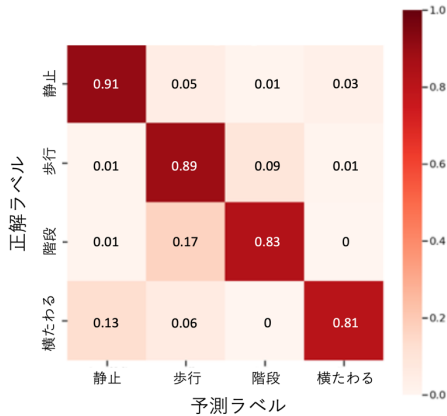


図 3 CNN-LSTM の混同行列

3.2 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、利用者の状態変化の有無の検知を行う。本稿では変化点検知手法の一例として、ChangeFinder [13]を利用する。また、変化点検知の閾値の設定には、ボリンジャーバンド [14]を利用する。併せて、変化点検知の計算タイミングに特性変化点を利用することで、計算回数を抑える手法を提案する。

3.2.1 特性変化点

変化点検知を行う際は、得られたデータを逐次処理するオンライン学習の手法が一般的である。提案手法では、Web ブラウザ内で行うことを想定しているが、状態区間推定に関する処理を加速度データの取得毎、すなわち数百ミリ秒毎に行うこととした場合、計算回数が多くなることが想定される。そのため、状態が変化した可能性がある時点であらかじめ検出し、その時点で状態区間推定に関する処理を行うこととする。

本研究では、状態が変化した可能性がある時点を「特性変化点」と定義し、検出にはモース理論を用いる。モース理論は、位相幾何学における理論であり、多様体の特性が変わる点を「臨界点」として定義することで、図形的特徴を捉えるものである。取得された加速度グラフから算出される臨界点、すなわち微分値が 0 の時点が特性変化点として定義し、この臨界点毎に状態区間推定の計算を行う。

3.2.2 変化点検知

ChangeFinder は、2 段階に渡って時系列モデルを学習させることにより、時系列データの変化点をスコアとして算出する手法である [13]。時系列データの変化が大きいかほど算出されるスコアも大きくなる。ChangeFinder の仕組みについて、以下の Step1-Step3 で説明する。

Step1: 時系列モデルの忘却学習と外れ値のスコアリング (第 1 段階学習)

時系列データの確率モデルとして、AR モデルを用いる。これをオンライン忘却型学習アルゴリズムである SDAR モデルによって学習し、得られた確率密度関数の列を $\{p_t(x): t = 1, 2, \dots\}$ とする。このとき、 $p_{t-1}(x)$ は $x^{t-1} = x_1, \dots, x_{t-1}$ から学習された確率密度関数である。各時点 t のデータ x_t の外れ値スコアを、以下の対数損失で計算する。

$$Score(x_t) = -\log p_{t-1}(x_t) \quad (1)$$

Step2: 外れ値スコアの平滑化

T を与えられた正数として、幅が T の平滑化区間を設ける。この平滑化区間 T によって、Step1 で得られた外れ値スコアの移動平均値 $\{y_t: t = 1, 2, \dots\}$ を計算する。スコアの移動平均値は以下のように計算される。

$$y_t = \frac{1}{T} \sum_{i=t-T+1}^t Score(x_i) \quad (2)$$

Step3: 時系列モデルの忘却学習と変化点のスコアリング (第 2 段階学習)

Step2 によって得られた新たな時系列データ $\{y_t: t = 1, 2, \dots\}$ に対して、再び AR モデルによるモデル化と SDAR モデルによる学習を行う。これによって得られる確率モデルの時系列を $\{q_t: t = 1, 2, \dots\}$ とする。さらに、 T' を与えられた正数として、時刻 t における T' -平均スコアを、以下の対数損失で計算する。

$$Score(t) = \frac{1}{T'} \sum_{i=t-T'+1}^t (-\log q_{i-1}(y_i)) \quad (3)$$

前述の Step1 の計算段階では、時系列データ中の外れ値のみが検出されるが、Step2 での外れ値スコアの平滑化を行うことでノイズに反応した外れ値が除去され、Step3 での 2 段階目の学習によって本質的な変動のみが検出される仕組みとなる。

3.2.3 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、時系列データの異常な値を捉える手法である [14]。時系列データの収まりやすいバンド幅を W 、平滑化区間 x の移動平均の値を \bar{x} 、標準偏差の値を σ 、 k を 1.0 以上 3.0 以下の任意の少数とした場合、 W は式 (4) で表される。

$$W = \bar{x} \pm k\sigma \quad (4)$$

このとき、 W に収まらない値を異常な値とする。また、データが正規分布に従うとしたとき、「 $\bar{x} \pm 1.0\sigma$ 」の範囲の中に収まる確率が約 68.3%、「 $\bar{x} \pm 2.0\sigma$ 」の範囲の中に収まる確率が約 95.4%、「 $\bar{x} \pm 3.0\sigma$ 」の範囲の中に収まる確率が約 99.7%となる。

4. 検証実験

4.1 実験概要

本提案における状態区間の推定手法の有効性を検証するための実験を、以下の手順で進める。

① データセットの収集

3 種類のシナリオに沿って行動した場合の加速度データおよび動作推定結果を収集する。行動するシナリオのラベルを実験者に入力してもらうことで、状態区間の正解ラベル (以降、正解区間) も合わせて収集する。

② 特性変化点の実験

①によって収集された加速度データと 3.1 節で述べた手法を利用して、特性変化点の検出を行う。また、加速度データの取得回数と特性変化点の検出数を比較し、全体の計算回数をどの程度抑制できるかを検証する。

③ 状態区間推定の実験

①によって収集された加速度データと 3.2 節で述べた手法を利用して、状態区間推定 (以降、推定区間) の決定を

行う。これにより、提案手法による推定区間の有効性を検証する。

④ パラメータの調整

①によって収集された正解区間と、③によって決定された推定区間とを比較し、区間数の差 N_{diff} が0の場合に区間時間の差 T_{diff} が最小になるパラメータを調査する。ここで N_{diff} とは、正解区間の数 $N_{correct}$ と推定区間の数 $N_{estimation}$ との差を示し、式(5)で表される。

$$N_{diff} = N_{correct} - N_{estimation} \quad (5)$$

また、 T_{diff} は、正解区間の開始時間 $Time_{correct}$ と推定区間の開始時間 $Time_{estimation}$ との差を示し、式(6)で表される。

$$Time_{diff} = Time_{correct} - Time_{estimation} \quad (6)$$

4.2 データセットの収集

4.2.1 データセット収集用アプリケーション

状態区間推定を行うために、加速度データと動作推定結果を収集する。データセットの収集には HTML と JavaScript で作成した Web アプリケーションを用いる。このアプリによって、加速度、角速度、動作推定結果のデータを 200 秒ごとに収集する。また、アプリ上に YouTube のライブ動画を設置しており、ユーザが休憩時などに動画を閲覧できるようになっている。上記の機能に加えて、アプリケーション使用時の場面ラベルを利用者に入力してもらうことで、そのときの状態を取得することを可能としている。入力された場面ラベルは、状態の正解ラベルとして利用する。

4.2.2 データ収集シナリオ

以下に、データ収集時の条件や実験シナリオをまとめる。

- 実験者：22 歳の男子大学生（以降、ユーザ A）、20 歳の女子大学生（以降、ユーザ B）、20 歳の男子大学生（以降、ユーザ C）の計 3 名
- 場所：岩手県立大学の構内
- 実験シナリオ：以下の 3 種類
 - (A) 大学の駐車場から歩いて校舎へ移動→途中から小走りで移動→校舎に着き次第、歩いて校舎内の教室へ移動→教室の椅子に着席（以降、シナリオ A）
 - (B) 廊下から図書館へ移動→館内で本を探す→館内の席に座って本を読む→館内の書籍貸し出し機へ移動→書籍貸し出し機で本を借りる→退館（以降、シナリオ B）
 - (C) 購買店舗前のベンチで休憩→購買店舗へ移動→購買店舗で買い物→ベンチへ移動→ベンチで休憩（以降、シナリオ C）

実験者に上記のシナリオで行動してもらい、行動時の場面が変わり次第、そのときの場面ラベルをアプリ上に入力することとする。なお、別の実験者によって、収集時の様子の動画を撮影することとする。

また、シナリオ A においては、校舎に向かって歩く距離や小走りを始めるタイミングについて事前に指示を行う。これにより、ユーザごとの状態に差は大きく生じてないことが考えられる。対して、シナリオ B, C においては、本もしくは商品を探す際の行動経路などは指示しないこととする。そのため、ユーザは自由に買い物などを行うことができるため、ユーザごとの状態に差が大きく生じることが

考えられる。そのため、シナリオ A を基本のシナリオとして定義し、シナリオ A に対する実験結果を中心に報告を行う。

4.2.3 収集されたデータセット

前述の条件に基づいて収集されたデータセットの例として、シナリオ A で行動したユーザ A によって収集された加速度の値を図 4 に示す。図中のマーカーの色と種類は動作推定ラベルを表し、青の実線は正解区間を表す。この図より、行動毎に加速度データと動作推定結果、正解区間の収集が行えていることが確認した。また、他のシナリオや他のユーザにおいても、データの収集が正確に行っていたことが確認した。

4.3 特性変化点の実験

4.3.1 実験概要

4.2 節の実験によって収集された加速度データに対して、特性変化点の検出を行う。以降の分析は、データ分析のため収集データに基づき Python (JupyterLab) を用いて行う。これにより、特性変化点による計算回数の抑制の有効性を検証する。この検証では、加速度データの取得回数と特性変化点の検出個数を比較する。これにより、この 2 つの時点のどちらかにおいて推定区間決定に関する計算を行うとした場合、より計算回数を減らせる指標がどちらかを確認する。また、比較を行う際には移動平均法による平滑化を施したデータも比較対象とし、平滑化による特性変化点の検出回数への影響を確認する。

4.3.2 計算回数の抑制に関する実験結果

シナリオ A で行動したユーザ A によって収集された加速度データに対し、特性変化点を検出した結果を図 5 に示す。加速度データに対してモース理論を適用することが可能であることが確認できる。また、他のシナリオや他のユーザにおけるデータに対しても、同様に特性変化点が検出されていることを確認した。

また、シナリオ A で行動したユーザ A によって収集された加速度データの取得回数と加速度データに対する特性変化点の検出回数、および移動平均法によって平滑化した加速度データに対する特性変化点の検出回数を図 6 に示す。移動平均法の平滑化区間は 5 で設定する。これにより、加速度データの取得タイミングではなく、特性変化点の取得タイミングを状態推定区間の計算タイミングに利用した場合、計算回数を約 39.5%減らせる結果となった。また、平滑化を施した加速度データに対する特性変化点についても、検出個数が減ることが分かった。この点から、特性変化点

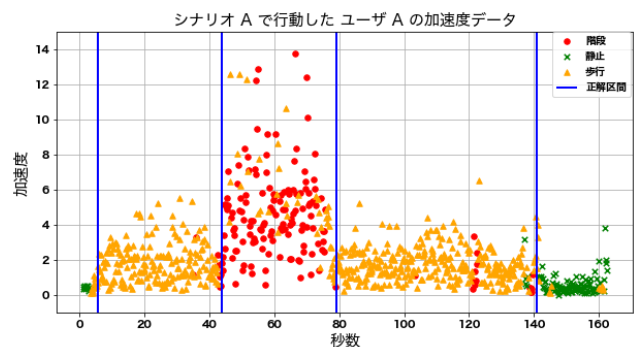


図 4 収集された加速度データ

と移動平均法によるデータの平滑化が、計算回数の抑制に繋がる結果となった。

4.4 状態区間推定の実験

4.2 節の実験によって収集された加速度データに対して、状態区間推定を行う。これにより、提案手法における ChangeFinder およびボリンジャーバンドの有効性を検証する。

4.4.1 パラメータの設定

状態区間の推定には、ChangeFinder, ボリンジャーバンドを用いる。なお、これらの手法で設定されるパラメータのうち、ChangeFinder の平滑化区間を 20, AR モデルの次数を 1, ボリンジャーバンドの平滑化区間を 40 に設定する。また、本稿において調整の対象としたパラメータを表 1 に示す。このうち、ChangeFinder の忘却率は、過去のデータの影響をどの程度反映するかを表すパラメータである。

なお、調整の指標には、式 (5), (6) で示した正解区間の個数と推定区間の個数の差 (以降、区間数の差) および正解区間の開始時間と推定区間の開始時間の時間差 (以降、区間時間の差) を利用する。本検証では、区間数の差が 1 つでも多い、もしくは少ない結果のパラメータについては、検証対象から外すこととする。よって、区間数の差が 0 かつ区間時間の差が最短であったパラメータを、そのデータに対する最適なパラメータとして利用することとする。

4.4.2 シナリオ A での結果

シナリオ A で行動したユーザ A によって収集された加速度データに対し、状態変化の検知時点を元に、推定区間を決定した結果を図 8 (a) に示す。正解区間を表す青の実線と、推定区間を表す赤の点線が概ね近い位置に図示されていることから、実験車によって付与された場面ラベルの切り替わりと近いタイミングで状態変化を検知できていることが分かる。ここでの正解区間と推定区間の時間のずれについては、移動平均法による値の平滑化からなる平均値の対応秒数のずれや、ChangeFinder によるスコアの平滑化からなるスコアの対応時点のずれが起因していると考えられる。また、ユーザ B, C においても、シナリオ A において収集されたデータでは、同様に正解区間と推定区間が近いタイミングで決定されることが確認できた。

4.4.3 シナリオごとの比較結果

シナリオ B, C で行動したユーザ A によって収集された加速度データに対し、状態変化の検知時点を元に、推定区間を決定した結果をそれぞれ図 8 (b), (c) に示す。この図から、正解区間と推定区間がある程度近い位置で図示されている。しかし、区間時間の差が約 30 秒の箇所が見られるなど、4.4.2 節の結果よりも正解区間と推定区間のずれが大きくなる結果となった。ここから、行動の自由度が低いシナリオ A に対して、行動の自由度の高いシナリオ B, C においては、状態区間の決定の精度が低くなる傾向があることが分かる。この点については、シナリオ B, C におけるユーザ B, C による収集データでも、同様に区間時間の差が大きくなる傾向が確認された。

また、図 8 (c) から、シナリオ C における 4 つ目の状態区間の開始時間 (図 8 (c) 中 4 本目の青の実線と 4 本目の赤の点線) が大きくずれていることが分かる。シナリオ C においてデータを収集している様子を撮影した動画におい

て、4 本目の赤の点線が示されている 120 秒付近を確認したところ、ユーザ A は購買店舗でレジ待ちをしていた。本来、4 つ目の正解区間は「ベンチへ移動」というラベルを付与していた。しかし、この結果から、本提案手法によって「レジ待ちをする」という、当初設定していなかった状態が推定されたことが分かった。このことから、人間の主観によって付与された状態のラベルと変化点検知によって機械的に付与された状態のラベルとでは、状態ラベルの内容に差異が生じるということが考えられる。この点については、シナリオ A におけるユーザ B, C による収集データでも、レジ待ちの行動の付近で新たな状態推定が行われていることが同様に確認された。

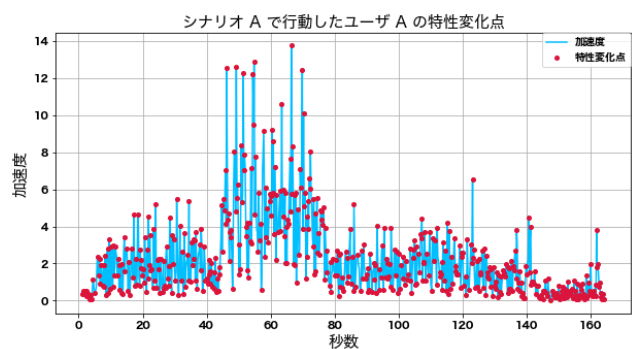


図 5 加速度データと検出された特性変化点

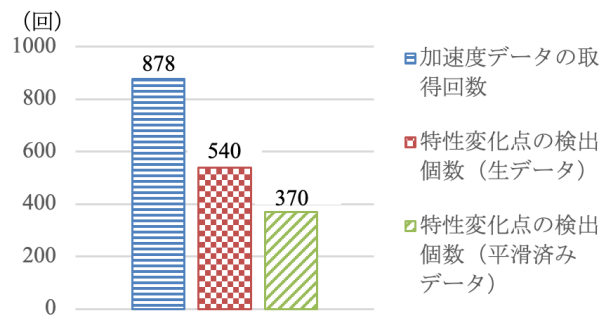


図 6 加速度データの取得回数と特性変化点の検出数

表 1 調整パラメータの概要

アルゴリズム名	パラメータ名	調整範囲
ChangeFinder	忘却率	0.01~0.99 (0.01 ずつ調整)
	ボリンジャーバンド	バンド幅
移動平均	平滑化区間	2.0~19.0 (1.0 ずつ調整)

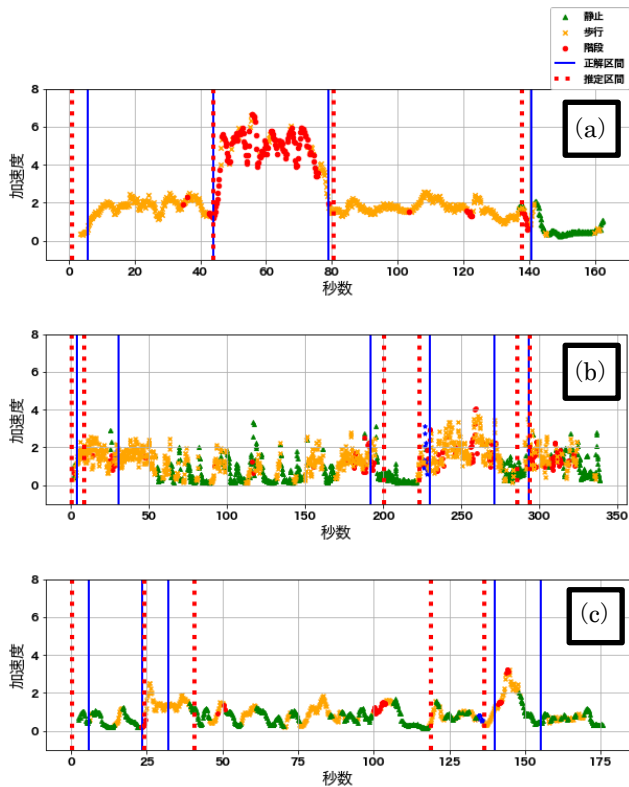


図 8 状態推定手法の概要図

表 1 最適パラメータ

		ユーザA	ユーザB	ユーザC
シナリオ A	忘却率	0.54	0.86	0.73
	バンド幅	1.97	1.55	1.57
	平滑化区間	12	19	19
シナリオ B	忘却率	0.14	0.46	0.66
	バンド幅	1.51	1.9	2.01
	平滑化区間	3	16	14
シナリオ C	忘却率	0.98	0.61	0.07
	バンド幅	1.04	1.45	1.65
	平滑化区間	9	5	4

4.4.4 パラメータ調整の検証結果

3つのシナリオにおける3人の実験者からのデータに対して、パラメータを調整した結果を表2に示す。ここから、シナリオB、Cについて、ユーザごとの平滑化区間や忘却率の値に差が生じているのが分かる。これに対し、シナリオAについては、シナリオB、Cほどの差が生じていないことが分かった。このことから、行動の自由度が低いシナリオであるほど、パラメータの差異が生じにくいことが分かった。

また、バンド幅のほとんどの値が、1.4から2.0の範囲に収まっている。ここでのバンド幅は変化点検知を行う際の閾値の範囲を示しているため、ユーザやシナリオの違いから生じるデータの分布の違いによる影響をあまり受けないということが言える。すなわち、今回の実験においては、1.4から2.0の範囲においてバンド幅の値を設定することがある程度適切であることが分かった。

5. おわりに

本稿では、モバイル端末利用者の状態を考慮したWebパーソナライズの実現を目的に、Webブラウザから取得されるセンサデータを利用した状態推定手法を提案した。モース理論やChangeFinder、ボリンジャーバンドの手法を利用することで、特性変化点や状態推定区間の検出が可能であることを示した。

今後は、本稿の提案内容の機能を有したWebアプリの実装、3章で述べた状態推定を行う機械学習モデルの構築などを行っていく。

参考文献

- [1] V. Salnen, H. Karjalainen, "Web personalization: The state of the art and future avenues for research and practice," *Telematics and Informatics*, vol.33, no.4, pp. 1088-1104, 2016年1月26日.
- [2] 日本経済新聞, "ローソン、食品ロス半減へ値引き情報位置データ活用," 2020年10月17日, <https://www.nikkei.com/article/DGXMZO65072170W0A011C2XQH000/>. [参照 2022年6月8日]
- [3] 小野寺斗弥, "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年.
- [4] L. Wu, H. Xiangnan, "A Survey on Accuracy-oriented Neural Recommendation: From Collaborative Filtering to Information-rich Recommendation," *IEEE Transactions on Knowledge and Data Engineering*, 2021年.
- [5] H. Ko, S. Lee, "A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields," *Electronics 2022*, 2022年.
- [6] web.dev, "What is FLoC?," 2021年3月30日. <https://web.dev/floc/>. [参照 2022年6月8日]
- [7] Google Japan Blog, "プライバシーサンドボックスの新しいTopics APIについて," Google, 2022年1月26日. <https://japan.googleblog.com/2022/01/topics-api.html>. [参照 2022年6月8日]
- [8] T. Okoshi, "Reducing Users' Perceived Mental Effort due to Interruptive Notifications in Multi-Device Mobile Environments," *UBICOMP '15*, 2015年.
- [9] R. Wampfler, "Affective State Prediction from Smartphone Touch and Sensor Data in the Wild," 2022年.
- [10] 大竹航平, "行動の文脈を考慮した状況のモデル化," 第9回日本感性工学会春季大会, 2014年.
- [11] G. Atluri, "Spatio-Temporal Data Mining: A Survey of Problems and Methods," *ACM Computing Surveys*, vol.51, no.4, pp.1-41, 2019年.
- [12] D. Miralles, "El Niño-La Niña cycle and recent trends in continental evaporation," *Nature Climate Change*, no.4, pp.122-126, 2013年.
- [13] K. Yamanishi, "A unifying framework for detecting outliers and change points from non-stationary time series data," *KDD '02*, 2002年.
- [14] NPO 法人日本テクニカルアナリスト協会, "ボリンジャーバンド," 2001年. https://www.ntaa.or.jp/technical/momentum/momentum_bollinger. [参照 2022年6月8日]