

# 時空間的なスマートフォンログ分析に基づく 利用者のストレス推定手法

濱谷 尚志<sup>1,a)</sup> 落合 桂一<sup>1</sup> 山本 直樹<sup>1</sup> 深澤 佑介<sup>1</sup> 木本 勝敏<sup>1</sup> 上西 康平<sup>2</sup> 太田 順<sup>2</sup>  
寺澤 悠理<sup>3</sup> 沖村 宰<sup>3</sup> 前田 貴記<sup>3</sup>

## 概要:

近年わが国ではスマートフォンの普及が進んでおり、スマートフォンを用いたヘルスケア技術に期待が高まってきている。スマートフォンはユーザが日々持ち歩き、使われるものであるため、ユーザの心理状態を反映することが期待できる。本研究では、ユーザのスマートフォン利用や持ち歩きによって収集されるログデータについて、時間的観点、空間的観点の両観点からログを分析し、利用者のストレスを推定する手法を提案する。その結果、20人のユーザにおいて、推定精度88.7%で利用者のストレスの高低度合いを推定可能であり、時間だけでなく空間的観点でログを集計することが有効であることを確認した。

キーワード: スマートフォン、ストレス、メンタルヘルスケア、滞留点抽出

## 1. はじめに

世界保健機関 (WHO) の調査によると、世界中で3億人もの人がメンタルヘルスの不調を抱えており、メンタルヘルスの不調による労働生産性の低下にともなう経済損失は1兆ドルにのぼるとも試算されている<sup>\*1</sup>。メンタルヘルスの不調に至る原因としては、性格や信条などの内的要因の他に、社会、文化などの外的要因が存在することが知られている。外的要因の中でも特に知られているのがストレスであり、メンタルヘルスの不調を防ぐためには、過度なストレスを避け、自身が受けているストレスを自覚し予防に努めることが重要である。

一方、昨今スマートフォンの普及率は世界的に増大しており、スマートフォンを活用したエージェントサービスやヘルスケア技術への関心が高まってきている。スマートフォンは複数のセンサを搭載していることから、歩数や移動経路の記録など、ヘルスケアへの応用がなされている。また、ユーザが多数のアプリや機能を使うことで様々なログが取得可能であり、ユーザの趣味趣向などを推定することが可能になりつつある。

このような背景を踏まえ、これまで本研究チームでは、スマートフォンの使い方を決まった時間単位で集計することにより、スマートフォンの持ち歩き方、使い方を特徴量化し、ウェアラブル心電計により取得したストレスの指標と照らし合わせることでユーザのストレスを推定する手法の研究を行ってきた。この方法により、どの時間帯にどのようなスマートフォンの使い方をしていくかといった時間的なコンテキストを考慮できる。一方で、決まった時間単位でのスマートフォンログ集計では、ログが発生した空間的なコンテキストを考慮できない可能性がある。例えば、同じ時間に電話をしている場合においても、自宅で電話をするのと職場で電話をするのとでは、電話の目的や電話によって生じる心理的な影響に差異があると考えられる。従来の研究でも、ユーザのメンタルヘルスの状態とユーザの位置情報に着目した研究がなされており、位置情報はユーザの心理状態を認識するために有用であると考えられている [1], [2]。しかしながら、位置情報はユーザにとってプライバシーに関わる情報でもあるため、必ずしも取得できるとは限らないという問題がある。

そこで本報告では、ユーザが訪れる場所によって、スマートフォンの利用ログの意味合いが異なりうるという想定のもと、ユーザの滞留地点に着目し、空間的にスマートフォン利用ログを分析することで、ストレス推定の精度向上を図る。さらに、空間的なログ分析において、ユーザのプライバシー配慮への必要性を鑑み、空間的にログを集計

<sup>1</sup> 株式会社 NTT ドコモ (NTT DOCOMO, INC.)

<sup>2</sup> 東京大学 (The University of Tokyo)

<sup>3</sup> 慶應義塾大学 (Keio University)

a) takashi.hamatani.rt@nttdocomo.com

\*1 [http://www.who.int/mental\\_health/world-mental-health-day/2017/en/](http://www.who.int/mental_health/world-mental-health-day/2017/en/)

するための滞留地点の検出において、位置情報を用いる方法だけでなく、位置情報を周辺 WiFi 情報で代替する方式を提案する。

実際に 20 人の合計 200 日以上ログデータにおいて、位置情報に基づきユーザの訪れた点を時空間的にクラスタリングした上で、寝食を行う地点と日中主に活動する地点、およびそれらに類する地点と、その他の滞在地点ごとにログデータを集計することでユーザの行動を空間的観点で特徴量化した結果、従来の時間的観点の特徴量のみを用いる場合と比較して、ストレスの推定精度が向上することを確認した。従来研究、ならびに著者らの先行研究に対する本報告の貢献は次の通りである。

- ユーザの滞留位置毎にスマートフォンの各種ログを特徴量化し、時間毎に抽出した特徴量と組み合わせることでユーザのストレスの高低推定精度 88.7%を確認し、時空間的コンテキストを踏まえた特徴量の有効性を確認した。
- プライバシーへの配慮のため、位置情報の代替として周辺 WiFi 情報を用いて重要な滞留点を抽出する方式を実装し、位置情報および WiFi 情報をそれぞれストレス推定に用いる場合の精度の比較を行った結果、同等の精度でストレスを推定できることを確認した。

## 2. 関連研究

### 2.1 スマートフォンを用いた人のストレス推定

スマートフォンには制御のために加速度センサや GPS など多数のセンサが搭載されており、ユーザの行動を受動的に取得することができる。こういったスマートフォンの特徴に着目し、これらのセンサから得られる情報を用いることでユーザのストレスを推定する研究が行われている [3], [4]。これらの研究では、スマートフォンの持ち歩きや利用傾向からユーザの感じているストレスを推定できる可能性を示している。また、Canzian・Musolesi の研究 [5] では、位置情報を用いて得られたユーザの移動軌跡を用いてユーザの抑うつ状態を推定する手法を提案している。また、Ware らは上記の研究を発展させ、大学生を対象に、大学の IT サービスへの接続情報（各 WiFi AP の MAC アドレス、接続デバイスの MAC アドレス、接続開始時間および接続時間）を用いることで、位置情報を用いず、アクセスポイント単位での位置情報を用いて抑うつ状態の推定を行っている [6]。以上の研究のように、スマートフォンから得られる情報を活用したユーザの心理状態推定の取り組みが数多く行われてきている。さらに、後者の研究のように、人が滞在する場所と人の心理状態の関係性に着目した研究も数多く行われており [1], [2]、本報告においても、滞在地点単位でスマートフォンのログを分析することで、ストレスの推定精度の向上を図る。

### 2.2 本研究チームのこれまでの取り組み

上述の研究など、従来のストレスの推定に関する研究では、PHQ-9[7]（生活に関する 9 個の質問）などのアンケートを用いてユーザの抑うつ度を計る方法が主流であった。こういったアンケートを用いた方法では、ユーザの生活の中で、定められたアンケートに回答してもらうだけでユーザの心理状態に関する正解データを得られる利点がある。一方で、回答結果がユーザの主観に依存するため、利用者が気づいていない潜在的な心理状態の抽出や、連続的に正解データを取得することが困難である。

そこで、筆者らの研究チームではウェアラブル心電計を用いて取得可能な LF・HF 比（ストレス状態と関連する指標）を正解データとし、スマートフォンより得られるログと組み合わせることで推定モデルを構築することにより、客観的なストレスの推定手法の研究を行ってきた。スマートフォンログを用いたストレスの推定可能性検討のため、11 種のスマートフォンログ（加速度、角速度、気圧、位置情報、充電、イヤホンの接続、画面のオンオフ、アプリの起動ログ、バッテリー、Google Activity Recognition、ストレージ容量）について、合計 128 次元の特徴量を実装し、推定モデルを構築することで 71.1%の精度でストレスを推定可能であることを示した [8]。さらに、ストレスの推定精度の向上のためストレスと相互作用があると考えられるユーザの注意機能に着目し、ストレスと注意機能の相互作用を考慮したストレス推定手法を提案し、実際に相互作用の考慮によりストレス推定精度が向上することを確認した [9]。

本報告では、従来の時間的なコンテキストを考慮した特徴量に加え、空間的なコンテキストを考慮した特徴量の抽出によりストレス推定精度の向上を図る。これまでに、重要な位置での SMS や電話の利用を特徴量として利用する研究はなされているが [10]、本研究ではアプリケーションの利用以外にも人の動きや環境などが空間的コンテキストと関連していると考え、重要な地点におけるスマートフォンログを網羅的に特徴量化し、ストレス推定への効果検証を行う。

## 3. ストレス推定手法

本研究では、著者らの先行研究で提案したストレス推定手法と同様の枠組みを用いてストレスの推定を行う。この際に、以降の 4 章で提案する空間的なスマートフォンログ集計に基づく新規特徴量を追加し、場所ごとのスマートフォン利用がユーザの心理状態と関連するかを明らかにすることが本報告の狙いである。本章ではストレス推定手法の概要を述べるが、詳細については先行研究 [9] を参照されたい。

### 3.1 概要

ストレス推定の流れを図 1 に示す。まず、スマートフォ

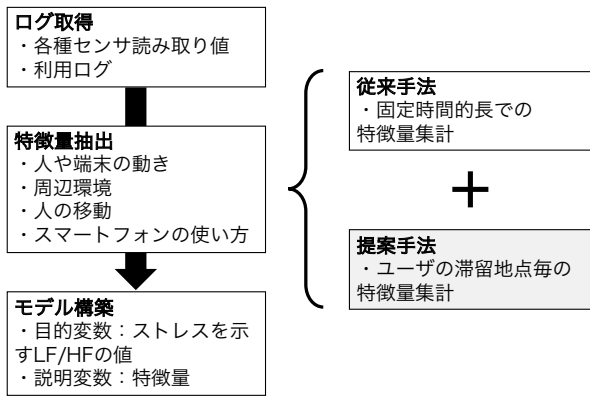


図 1 スマートフォンログを用いたストレス推定の概要

表 1 収集するスマートフォンログ

ログ種別	取得頻度
加速度 (3 軸), 角速度 (3 軸), 端末傾き (3 軸)	1 秒毎
気圧, 周辺照度, バッテリ残量, 接続先 WiFi アクセスポイント情報, 周辺 WiFi アクセスポイント情報, Google Activity Recognition 結果, 位置情報	5 分毎
ストレージ空き容量, 曜日情報	1 日 1 回
画面のオン・オフ, アプリの起動・終了, イヤホンジャックへの接続, 充電器への接続	イベント発生時

ンより取得可能なログを収集し、ユーザのスマートフォン所持や利用の特徴を表す特徴量を抽出する。抽出した特徴量は、ストレス状態の正解データとなる LF・HF 比のデータと組み合わせ、教師あり機械学習により特徴量を入力としたストレス推定モデルを構築する。モデル構築後、モデルの推定精度の評価を行う場合は、抽出した特徴量をモデルに与えることで得られる推定結果と、正解データを比較することにより実施する。

### 3.2 スマートフォンログの収集

近年、日本で利用されているスマートフォンの OS のシェアは Android OS および iOS が主である。両者を比較すると、Android OS ではアプリケーションの利用履歴などが取得できる一方で、iOS ではアプリ利用履歴へのアクセスが制限されており、Android OS の方がより多くのログを取得可能である。したがって本研究では、スマートフォンを用いたストレスの推定可能性の検討のため Android OS を対象とし表 1 に示すログを取得する。スマートフォンの状態に関するログ（センサ類や位置情報など）の取得間隔は小さければ小さいほどより詳細な情報が得られるが、取得データ数の増加に伴いバッテリー消費量やログファイルのサイズが増大するため、トレードオフの関係になる。本研究では上記の関係を鑑み、表 1 に定める間隔でログを取得する。

### 3.3 特徴量の抽出

上述のスマートフォンログでは、取得タイミングごとの

表 2 生成する特徴量

種別	特徴量
ユーザやスマートフォン動き	加速度・角速度・端末傾きにおける平均値・標準偏差・最大値・最小値・最大値と最小値の差・ノルム・相関係数, Google Activity Recognition における各 Activity クラスの割合
周辺環境	気圧と照度の平均値・標準偏差・最大値・最小値・最大値と最小値の差, 接続および観測した WiFi アクセスポイントの数
ユーザの移動	緯度・経度・高度（初期位置を基準とする相対座標系）の最大値・最小値・最大値と最小値の差, 先行研究 [5] で用いられた特徴量（移動距離, 移動距離の分散, 重要地点数など）
スマートフォンの使い方	画面オン・オフの回数・オンの時間平均, 偏差, 最大値・最小値, 各カテゴリごとのアプリの起動回数・起動時間の平均, 偏差, 最大値・最小値, アプリ（各カテゴリ）起動, および画面オンにおける一時間ごとの確率分布の Bhattacharyya 係数 [11]（対前日比, 対前週比, 対ユーザ平均比）

値が記録されているが、ログそのものは膨大な数値の羅列であるため、ログそのものを用いてストレスの状態を推定することは困難である。本研究ではユーザがストレスを受けているときに表れる行動をスマートフォンログより検出することを目的とし、以下の 4 種類の観点でログより特徴量を抽出する。(1) 人やスマートフォンの動きに関する特徴量 - ストレスを受けることによりユーザやスマートフォンの動きに影響を与える。(2) 周辺環境に関する特徴量 - 周辺環境の変化がストレスに影響を与え、環境の変化でストレスによる人の行動を解釈する。(3) 人の移動に関する特徴量 - ストレスの状態が人の移動方法や訪れる場所に影響を与え、訪れる場所により人のストレス状態が変化する。(4) スマートフォンの使い方に関する特徴量 - ストレスの状態によりスマートフォンを使う目的や使い方が変化する。

著者らのこれまで研究では、上記の特徴量を 1 時間毎、6 時間毎および 24 時間毎に抽出し、ストレス推定のための特徴量としていた。一方、本報告では上記特徴量を決まった時間単位ではなく、ユーザが訪れた場所毎に集計することで空間的な観点からスマートフォンログを集計しストレス推定に活用する。すなわち、抽出する特徴量の種類自体はこれまでの手法と同様であるが、決まった時間単位ではなくある地点に滞在した時間帯のログを対象として特徴量を抽出する。特徴量の詳細を表 2 に示す。各特徴量の詳細な抽出方法は先行研究 [9] を参照されたい。

### 3.4 教師データの作成

2.2 節で述べた通り、本研究ではユーザの客観的なストレスの推定のため、胸部心電計を用いて計測した心電波形より得られる LF・HF 比 [12] をストレスの正解データとして収集する。LF および HF はそれぞれ心電波形の低周波

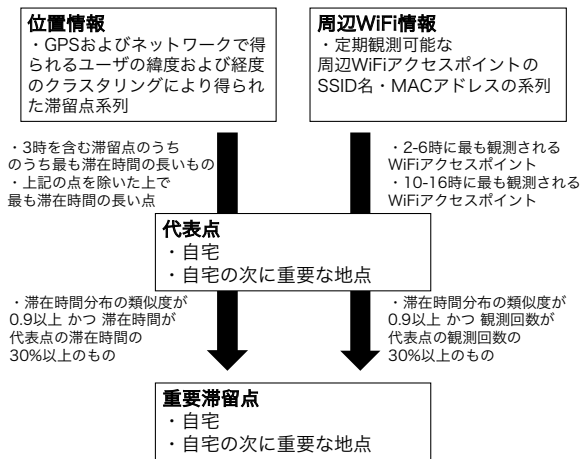


図 2 位置情報または WiFi 情報を用いたユーザの重要地点抽出アルゴリズム

成分と高周波成分を表す。LF・HF 比は交感神経と副交感神経の活性度のバランスを示すことが知られており、ユーザがストレス状態にあると LF・HF 比の値が大きくなることが知られている。本研究では大局的なユーザのストレスを推定することを目的としており、心電計より得られた 5 分毎の心電波形に対し LF・HF 比を計算し、13 から 16 時までの LF・HF 比の値を平均化することで一日の LF・HF 比の代表値を得る。LF・HF 比は人によって異なりうる値であるため、ユーザが受けているストレスが高いか否かの判断のため、ある LF・HF 比の値がユーザの平均値より高いかどうかに着目する。この方法により、ユーザの一日の LF・HF 比の代表値を 1 (ストレスが高い) または 0 (ストレスが低い) に変換し、後述の推定モデル構築における教師データとする。

### 3.5 推定モデルの構築

3.3 節で述べたスマートフォンログより得られた特徴量を説明変数、3.4 節で作成した教師データを目的変数とする教師あり機械学習により、ストレス推定モデルを構築する。機械学習のアルゴリズムは複数手法の比較の結果 extreme gradient boosting (XGBoost) [13] を用いる。学習において説明変数の次元が膨大となるが、膨大な次元数での学習は次元の呪いにより学習が困難になるため、XGBoost のモデル学習時に得られる各特徴量の重要度の値に基づき重要な特徴量のみを選択することで次元数を削減する。

## 4. 提案手法: 重要滞留点におけるスマートフォンログの特徴量化

### 4.1 重要滞留点抽出方式

ユーザの位置情報を表すログを分析することで、ユーザがどういった場所を移動したかや、滞在したかを解釈することが可能である。ユーザのストレスと滞留する地点の関係の分析のためには、単なる物理的な滞留位置 (緯度、経

度など) ではなく、滞留地点の意味 (自宅や職場などの論理的な位置) を解釈することが重要である。また、論理的な位置情報を解釈することで、異なるユーザの滞留点情報の比較が可能になる。

そこで本報告では、図 2 に示すアルゴリズムを用いてユーザにとって重要な滞留点を抽出する。本研究では人間の生活において、寝食と活動の二種類を行う場所がユーザにとって特に重要であると想定し、第一重要地点 (主に寝食をする場所)、および第二重要地点 (主に活動をする場所) の二地点を推定し、それらの地点におけるスマートフォンログより特徴量を抽出する。重要地点の数を限定することで、数多くのユーザから共通して特徴量を抽出することができるため、異なるユーザを比較する推定モデルを構築可能であると考えられる。さらに、第一重要地点および第二重要地点に滞在していない時間帯におけるログからも特徴量を抽出する。以上の 3 種類の論理的滞留位置を位置情報または周辺 WiFi アクセスポイント情報を用いて抽出する方式を提案する。

位置情報は主に GPS を利用するため、緯度経度単位での位置情報の取得が可能であり高精度な位置情報が得られる。しかしながら、継続的な取得においてはバッテリー消費量の増大が問題となる。一方で、周辺 WiFi アクセスポイント情報は WiFi 接続のために受動的に収集している情報のため、WiFi 接続を利用する多くのスマートフォンにおいて追加のバッテリー消費を要しない利点がある。その一方で、得られる情報はアクセスポイントの名称や MAC アドレスであるため、物理的な位置情報を得ることができない欠点がある。以上の 2 種類のログは、それぞれユーザの論理的な位置を推定できる可能性があるが、性質が異なるため、本報告では 2 種類のそれぞれを用いたユーザの重要滞留点での滞在の抽出方式を検討する。

#### 4.1.1 位置情報を用いた重要滞留点抽出

位置情報は緯度経度の系列であるため、ユーザの行動理解に用いる場合地図情報などと組み合わせるの必要があり解釈が困難である。そのため、ユーザが類似する座標で滞留していた場合に滞留点としてまとめる処理を行い解釈を容易にする。本研究では西田らの時空間クラスタリング手法 [14] を利用することにより、地理的に近かつまとまった時間滞在している地点を滞留点として抽出する。

クラスタリングの結果得られた滞留点の系列からユーザの第一重要地点および第二重要地点を推定するため、提案手法では多くのユーザが就寝していると考えられる午前 3 時にユーザが最も滞在している滞留点を第一重要地点と推定する。また、就寝以外、すなわち主な活動時間帯においてユーザが滞在している地点をユーザにとって重要な活動地点であるとみなし、第一重要地点を除いた上で最も滞在時間の長い滞留点を第二重要地点と定める。以上の方法を 5.1 節で述べるデータセットに適用し得られた、全ユーザに

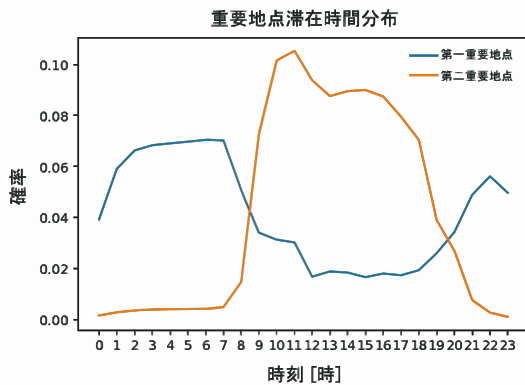


図 3 第一重要地点および第二重要地点への滞在時間帯分布推定結果（位置情報利用）

おける第一重要地点および第二重要地点への滞在時間帯分布の推定結果を図 3 に示す。以上の方法により、第一重要地点および第二重要地点を表す代表点をそれぞれ求める。

また、本研究では 5.1 節で述べるデータセットを用いてモデルの構築や評価を行うが、このデータセットは複数の勤務地を有する会社員を対象に収集したものである。したがって、各ユーザにとっての第二重要地点は主に勤務する地点を表すが、勤務地は一つとは限らないため、複数の勤務地をまとめて第二重要地点群として処理することが望ましい。同様に、第一重要地点として、自宅や長期出張の宿泊先など、寝食を行う地点が複数存在する可能性がある。

以上のように、第一（第二）重要地点の代表点の他に、同様の役割を持った滞留点をまとめて第一（第二）重要地点として扱えるようにする必要がある。滞留点の役割が第一（第二）重要地点の代表点と同様かどうかを判定するために、同じ目的で滞在する点と同様の時間分布で滞在するという想定のもと、滞在時間帯の分布の類似度が高い滞留点を同一目的の滞留点として扱う。具体的には、各重要地点に対し、一時間ごとの滞在確率分布の比較における Bhattacharyya 係数 [11] が 0.9 以上の滞留点を同一の役割を持つ滞留点とみなす。さらに、その滞留点が重要であるかどうかの判断のため、滞在確率分布の類似度が高いだけでなく滞在時間の長さにも注目する。具体的には、第一（第二）重要地点の代表点の滞在時間に対して、滞在時間が 30% 以上である滞留点を重要な滞留点であると判断する。以上の二種類の処理により、第一（第二）重要地点の代表点と同様の滞在傾向で、かつ一定の割合以上の滞在時間を持つ滞留点を重要滞留点群としてまとめる。以上の処理におけるパラメータ (0.9, 30%) は本研究で用いるデータセットに対して経験的に適切な値を一時的に設定しているが、本来はユーザの生活様態に応じて最適な値を設定することが望ましい。しかしながら、最適なパラメータの決定手法の確立には一層の検討を要する。

以上の方法により重要地点、および重要地点と同一の意味を持つ滞留点を推定するためにはある程度の期間の位置

表 3 周辺 WiFi 情報の例

時刻	SSID 名	MAC アドレス
2019/1/1 0:00:03	myaccesspoint-g	11:22:33:44:55:66
2019/1/1 0:00:03	otonari-home-wifi	44:55:66:77:88:99
2019/1/1 0:05:03	myaccesspoint-g	11:22:33:44:55:66
⋮	⋮	⋮
2019/1/1 12:00:03	office-wifi	66:55:44:33:22:11
⋮	⋮	⋮

情報ログを収集することが必要である。本報告は提案方式により生成した特徴量のストレス推定への効果検証を目的としており、重要地点の推定は行える前提で特徴量を生成するために、5.1 節で述べるデータセットの各ユーザにおいて収集した全ての位置情報ログデータを用いて重要地点の推定を行う。こうして得られた第一（第二）重要地点群を示す座標のリストと、各日の位置情報ログに対し時空間クラスタリング手法 [14] を適用して得られた滞留点の系列を照らし合わせ、いずれかの座標の近傍 (500 メートル以内) に滞在している場合、各重要地点に滞在していると判定する。

#### 4.1.2 周辺 WiFi 情報を用いた重要滞留点抽出

周辺 WiFi 情報は、表 3 の例で示すように観測した時間と SSID 名、MAC アドレスの組であり、空間情報は含まれていない。一方で、一度の観測で複数の SSID 名・MAC アドレスの組を取得可能であり、観測可能な WiFi アクセスポイントは空間によって異なるため、観測したアクセスポイントの集合は、それらのアクセスポイントを観測可能な地点、すなわち固有の位置を表す情報として利用することができる。

周辺 WiFi 情報を用いて第一重要地点と第二重要地点への滞在を推定するために、図 3 で示す第一重要地点および第二重要地点への滞在時間帯分布を用いる。図 3 を参考に、午前 2 時から午前 6 時に最も観測された WiFi アクセスポイントを第一重要地点を表すアクセスポイント情報とし、午前 10 時から午後 4 時に最も観測された WiFi アクセスポイントを第二重要地点を表す情報とする。以上のルールにより、周辺 WiFi を利用して推定した第一重要地点および第二重要地点への滞在時間帯分布を図 4 に示す。図 4 と図 3 のそれぞれの第一重要地点、および第二重要地点における Bhattacharyya 係数はそれぞれ 0.995, 0.993 であり、周辺 WiFi 情報を用いる場合においても、位置情報を用いる場合と同様の傾向で重要な位置への滞留を推定できることが分かる。

また、4.1.1 節同様に、第一重要地点および第二重要地点が複数地点存在することを考慮するために、各アクセスポイントに対し、アクセスポイントの検出頻度および検出される時間帯分布の類似度に基づいたフィルタを適用し、第一重要地点および第二重要地点と同様の滞留傾向を示すアクセスポイントかどうかを判断する。具体的には、第

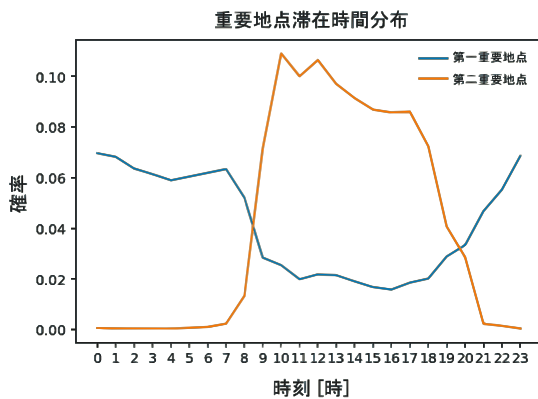


図 4 第一重要地点および第二重要地点への滞在時間帯分布推定結果 (周辺 WiFi 利用)

一 (第二) 重要地点の代表地点を表すアクセスポイントの検出回数に対し検出回数が 30%以上であるアクセスポイントで、かつ一時間ごとの滞在時間帯分布の比較における Bhattacharyya 係数が 0.9 以上のアクセスポイントを第一 (第二) 重要地点群としてまとめる。

以上の方法により重要地点、および重要地点と同一の意味を持つ滞留点を推定することが可能であるが、4.1.1 節と同様に、今回は 5.1 節で述べるデータセットの各ユーザにおいて収集した全ての周辺 WiFi ログデータを用いて重要地点の推定を行う。こうして得られた第一 (第二) 重要地点群を表すアクセスポイント情報のリストと、各日の周辺 WiFi ログを照らし合わせ、いずれか一つ以上のアクセスポイントを検出できている場合に各重要地点に滞在していると判定する。

## 4.2 重要滞留点における特徴量抽出方法

4.1.1 節および 4.1.2 節で述べた方式で抽出した各重要地点、および、その他の地点での滞在時間に対し、表 2 に示す特徴量の抽出を行う。提案手法で集計対象となるログは空間的に近いログを集計するため、表 2 のうち人の移動に関する特徴量は行わないものとする。以上の方法により、3 箇所 (第一重要地点、第二重要地点、その他) において各 195 種類、合計 585 個の特徴量を抽出する。

## 5. 性能評価

### 5.1 データセット

本報告では提案手法による新規特徴量のストレス推定への効果検証のため、著者らの先行研究で収集した株式会社 NTT ドコモの R&D 社員 34 人の合計 554 日のデータセットを用いる。ユーザにとって重要な滞留点でのスマートフォンログはユーザの心理状態を反映するという本研究における仮説の検証のため、性能評価では第一重要地点、および第二重要地点のいずれかまたは両方に滞在した日を対象としてストレス推定モデルを構築する。

上記データセットに対し、4.1.1 節で述べた位置情報に基

づく滞留点推定をユーザ毎に適用した結果、第一重要地点と第二重要地点への滞留を抽出できたログは 20 人の 212 日 (高ストレス: 84 日, 低ストレス: 128 日) であることを確認した。また、比較のため、4.1.1 節で述べた位置情報に基づく滞留点推定および 4.1.2 節で述べた周辺 WiFi 情報に基づく滞留点推定の両方で第一重要地点と第二重要地点への滞留を抽出できたログは 20 人の 174 日 (高ストレス: 69 日, 低ストレス: 105 日) であった。以降の 5.3 節、5.4 節の評価では、上記二種類のデータセットを用い、

- 本報告で提案する新特徴量の効果検証
- 本報告での新特徴量の生成方法 (位置情報または周辺 WiFi 情報を利用) の比較

の評価をそれぞれ行う。

### 5.2 評価手法

先行研究 [9] で導入した合計 7437 次元の特徴量、および本報告で提案する 585 次元の新特徴量を用いたストレス推定モデルの性能への影響評価のため、一人抜き交差検証により新規ユーザを想定したモデルの推定精度を求める。具体的には、20 人のユーザの 19 人のデータを用いてストレス推定モデルを構築し、残る 1 人のデータでモデルの性能評価を行うことを 20 通り全てのデータ分割方法で行い、20 人それぞれの性能評価結果を統合する。これにより、推定モデルの構築において評価対象ユーザのデータを全く用いず、未知のユーザを想定した評価を行う。

推定問題はストレスが高いか低いか (= 高ストレス状態が陽性か陰性か) の二値分類問題とし、推定精度の指標は、Accuracy (陽性と陰性を合わせた全体の正解率)、Precision (適合率: 陽性と検知したもののうち真に陽性である割合)、および Recall (再現率: 真の状態が陽性であるもののうち検出した割合) を用いる。

### 5.3 空間特徴量を用いたストレス推定精度

まず、先行研究で提案した時間的観点で生成した特徴量と、本報告で提案する空間的観点の特徴量の効果の比較、および併用した際の効果検証のため、二種の特徴量を独立して使用した場合、および併用した場合のストレス推定精度を表 4 に示す。この結果より、時間的特徴量と空間的特徴量を比較すると、時間的特徴量を用いたストレス推定モデルの方が推定精度が高いことが分かる。この理由として、時間的特徴量の次元数が多く、ストレス推定に有効な特徴量をより多く抽出できていたと考えられる。また、今回のデータセットでは同じ職種のユーザを対象にストレス推定を行っているため、生活の時間帯も概ね類似していたため、時間的観点での特徴量が有効であったと考えられる。

一方で、時間的特徴量と空間的特徴量を併用する場合、ストレスの推定精度が各指標を通じて向上しており、Accuracy が 88.7% を達成していることが確認できる。す

表 4 特徴量ごとのストレス推定精度 (ユーザ数 20, サンプル数 212)

用いる特徴量	特徴量数	Accuracy	Precision	Recall
時間的特徴量 (従来手法)	7437	81.6%	79.2%	72.6%
空間的特徴量 (提案手法)	585	70.3%	62.1%	64.3%
時間的特徴量+空間的特徴量 (位置情報ベース)	8022	88.7%	86.6%	84.5%

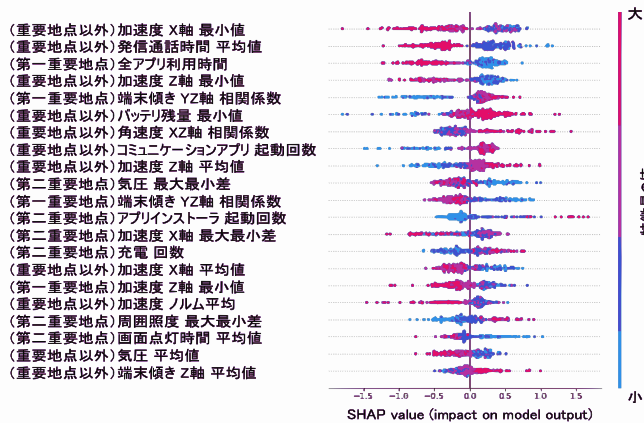


図 5 ストレス推定における空間的特徴量の影響

なわち、空間的特徴量単独だと時間的特徴量の推定性能には及ばないものの、両者を組み合わせることでストレス推定に有効なより多くのパターンを抽出できたことを示唆している。したがって、時間的観点に加え、空間的観点でスマートフォンログを分析することはユーザのストレス推定において有効であったといえる。

さらに、ストレス推定において本報告で提案した空間的観点で抽出した特徴量がどう作用したかを明らかにするため、SHAP value (Shapley additive explanations) を用い、各特徴量の推定への影響を算出する。図 5 に特に推定に寄与した特徴量の影響を可視化した結果を示す。図では縦方向に各特徴量名が、横方向に SHAP value が描画されている。図の各点は各推定サンプル (本評価ではある一日の特徴量の値を示す) を表しており、各点における色は特徴量の値の大小を表している。横軸 (SHAP value) は特徴量の寄与度を表しており、左側に位置するほど推定結果 (ストレスが高い確率) がマイナスになる方向に作用する。逆に、右側に位置するほどストレスが高い確率がプラスになる方向に作用する。したがって、各特徴量について、横軸に対する分布および点の色とを照らし合わせ、色が分かっている特徴量ほど、推定において有効な特徴量であることが分かる。

この結果より、各特徴量 (縦軸) が色ごとに左右に分かれて分布していることが確認できる。すなわち、特徴量の大小がストレスの高低推定において寄与していることを表しており、今回提案した重要滞在地点ごとに特徴量を生成するアプローチがストレス推定において有効であったといえる。具体的には、第一および第二重要地点だけでなく、その他の地点でのユーザの行動特徴量がストレス推定にお

表 5 異なる空間特徴量集計方法におけるストレス推定精度 (ユーザ数 20, サンプル数 174)

用いる特徴量	Accuracy	Precision	Recall
時間的特徴量 (従来手法)	77.6%	71.4%	72.5%
時間的特徴量+空間的特徴量 (位置情報ベース)	81.0%	75.7%	76.8%
時間的特徴量+空間的特徴量 (周辺 WiFi 情報ベース)	81.0%	79.0%	71.0%

いて重要であることが明らかになった。さらに、加速度では平均値や最小値が大きい (動きがある) 場合にストレスを低く推定する傾向が示されており、人がストレスを受けていることを、人やスマートフォンの動きセンサから検出できる可能性が示されている。また、重要地点以外での発信通話時間が長いとストレスを低く推定する傾向がある一方で、コミュニケーションアプリの起動回数は多いとストレスを高く推定する傾向があり、コミュニケーションに関しては方法や頻度の観点で、ストレスと複雑な関係を持っていることが示唆されている。加えて、第二重要地点では周囲の明るさや気圧などの環境条件に関する特徴量も反応しており、主に活動を行う場所における環境要素も人のストレスと関連していることが明らかになった。以上の分析を通じて、先行研究で提案した時間的観点での特徴量に加え、本報告で提案した空間的観点での特徴量集計方式によりストレスと関連するユーザの行動を捉える可能性がより高まったと考えられる。このことは、従来手法により生成した特徴量と本報告で提案した特徴量の組み合わせによりストレス推定精度が向上したことから裏付けられている。

#### 5.4 位置情報を周辺 WiFi 情報で代用した場合

前節の評価では、スマートフォンで取得した位置情報に基づいて生成した空間特徴量によるストレス推定への効果検証を行った。本節では、位置情報に基づき生成した空間特徴量、および周辺 WiFi 情報のみを用いて生成した空間特徴量のストレス推定における効果の差を検証する。

両者の比較のため、いずれの方式でも空間特徴量生成を実施できた 20 人ののべ 174 日のデータセットを用いてストレス推定の精度の評価を行った結果を表 5 に示す。前節の評価と同様に、ベースラインとして時間的特徴量のみを用いた場合と、時間的特徴量と位置情報を利用した空間特徴量を併用する場合、および時間的特徴量と周辺 WiFi 情報を利用した空間特徴量を併用する場合のストレス推定精度を示している。この結果から、全体的なストレス高低の成功率 (Accuracy) において、いずれの方式で生成した空

間特徴量においても、ストレス推定精度の向上に寄与したことが確認できる。

また、Precision と Recall に着目した場合、位置情報を利用した場合はベースラインに対して一意に性能が向上しており、陽性・陰性のいずれのサンプルにおいても推定精度の向上を確認できた。一方で、周辺 WiFi を用いた場合は高ストレス時における再現率 (recall) がベースラインに対して 1.5% 低下している。両者の傾向が異なる理由としては、重要地点に滞留しているかどうかの判定の粒度が影響していると考えられる。位置情報を用いた滞留点検出ではクラスタリングの半径を 500 メートルに設定しており、重要地点そのものだけでなく、近傍での滞在も含めて滞在と判定している。一方で、周辺 WiFi 情報を用いた滞留点検出では、重要地点固有の WiFi アクセスポイントを検出できない場合には滞在と判定しないため、WiFi 設備の配置状況によっては、部屋単位やフロア単位など、より細かい単位で重要地点かどうかの判定が行われる。したがって、実際に同じ座標に滞在していたとしても、重要地点かどうかの判定に差が生じる可能性があり、その結果、両者の特徴量が異なりストレス推定モデルの精度に影響したものと考えられる。しかしながら、いずれの特徴量を用いる場合も全体的なクラス分類精度としては向上しており、時間的観点だけでなく、空間的な観点でスマートフォンログを集約することはユーザの状況認識において有用であるといえる。

## 6. まとめと今後の課題

本報告では、ユーザが日々持ち歩くスマートフォンはユーザの心理状態を反映するとの想定のもと、スマートフォンから得られるログデータを用いて利用者のストレスを推定する手法を提案した。提案手法では、時間的観点でのログ集計だけでなく、空間的観点でのログ集計がユーザの心理状態をとらめるために有効であると想定し、位置情報、および周辺 WiFi の情報を用いて空間的観点でユーザの行動を表す特徴量を集計する方式を提案した。提案手法による新規特徴量の有効性検証のため、時間的観点で集計した特徴量と組み合わせてストレス推定モデルを構築、評価を行った結果、その結果、20 人のユーザにおいて、推定精度 88.7% で利用者のストレスの高低度合いを推定可能であることを確認し、時間だけでなく空間的観点でスマートフォンのログを集計することが有効であることを確認した。また、評価結果より、重要な滞在点でのスマートフォンログだけでなく、重要地点以外でのスマートフォンログ (動きのセンサやアプリの利用など) より得られた行動特徴量もストレスの推定に寄与することを確認した。

今回の評価では、従来研究で用いた特徴量を基準として本報告で提案した新しい特徴量の有効性の検証を行った。一方で、従来の特徴量は位置情報を用いて生成するものが

含まれているため、今後の課題として、プライバシーを考慮するために全ての特徴量で位置情報を用いない場合の精度検証を行う必要がある。

## 参考文献

- [1] Palmius, N. et al.: Detecting bipolar depression from geographic location data, *IEEE Trans. on Biomedical Engineering*, Vol. 64, No. 8, pp. 1761–1771 (2016).
- [2] Saeb, S. et al.: The relationship between mobile phone location sensor data and depressive symptom severity, *PeerJ*, Vol. 4, p. e2537 (2016).
- [3] Bogomolov, A. et al.: Daily stress recognition from mobile phone data, weather conditions and individual traits, *Proc. of the 22nd ACM international conference on Multimedia*, ACM, pp. 477–486 (2014).
- [4] Sano, A. and Picard, R. W.: Stress recognition using wearable sensors and mobile phones, *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, IEEE, pp. 671–676 (2013).
- [5] Canzian, L. and Musolesi, M.: Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis, *Proc. of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, ACM, pp. 1293–1304 (2015).
- [6] Ware, S. et al.: Large-scale Automatic Depression Screening Using Meta-data from WiFi Infrastructure, *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No. 4, p. 195 (2018).
- [7] Kroenke, K. et al.: The PHQ-9: validity of a brief depression severity measure, *Journal of general internal medicine*, Vol. 16, No. 9, pp. 606–613 (2001).
- [8] Yamamoto, N. et al.: Physiological Stress Level Estimation Based on Smartphone Logs, *Proc. of the 11th International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, IEEE (2018).
- [9] 濱谷尚志ほか: ストレスと注意機能の相互作用を考慮したスマートフォンログからのストレス推定手法, マルチメディア, 分散協調とモバイルシンポジウム 2019 論文集, Vol. 2019, pp. 759–769 (2019).
- [10] Boukhechba, M. et al.: Monitoring social anxiety from mobility and communication patterns, *Proc. of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ACM, pp. 749–753 (2017).
- [11] Cha, S.: Comprehensive survey on distance/similarity measures between probability density functions, *City*, Vol. 1, No. 2, p. 1 (2007).
- [12] Vollmer, M.: A robust, simple and reliable measure of heart rate variability using relative RR intervals, *2015 Computing in Cardiology Conference (CinC)*, IEEE, pp. 609–612 (2015).
- [13] Chen, T. and Guestrin, C.: Xgboost: A scalable tree boosting system, *Proc. of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, ACM, pp. 785–794 (2016).
- [14] Nishida, K. et al.: Probabilistic identification of visited point-of-interest for personalized automatic check-in, *Proc. of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, pp. 631–642 (2014).