

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

濱谷 尚志¹ 落合 桂一¹ 山本 直樹¹ 深澤 佑介¹ 木本 勝敏¹ 霧生 和樹² 上西 康平² 太田 順³
寺澤 悠理⁴ 沖村 宰⁵ 前田 貴記⁵

概要：精神疾患の患者数は年々増加傾向にあり，原因となるストレスの状態を把握し予防に努めることが重要である．本研究ではユーザが日々利用するスマートフォンに着目し，スマートフォンで取得可能なセンサ値および端末利用ログを用いてユーザのストレス状態を推定する手法を提案する．提案手法ではストレスと相互作用をもつ注意機能に着目し，注意機能の状態を考慮することでストレス推定精度の向上を図る．実際に 32 人の会社員から収集したのべ 554 日のデータセットを用いて評価を行った結果，注意機能を考慮せずにストレスを推定する方式と比較して，提案手法によりストレスの推定精度を平均 5.4% 高められることを確認し，ストレスの推定において注意機能との相互作用を考慮することが有効であることが分かった．

1. はじめに

近年わが国では精神疾患の患者数が年々増加傾向にあり，メンタルヘルスケアに関する社会的関心が高まってきている [1]．過度なストレスは精神疾患の発症原因となることが知られているため [2]，自身が受けている精神的なストレスを自覚し予防に努めることが重要である．

こういった社会背景の中で，スマートフォンなどの端末を用いてユーザの精神状態を推定する取り組みが数多くなされている [3]．スマートフォンは個々人が所持し，持ち歩き，かつ利用するものであり，その使われ方にユーザの精神状態が色濃く反映されるため，ユーザの精神状態を推定するのに有用である．これまでに本研究チームでもスマートフォンから取得可能なログデータを用いてユーザの精神的なストレスを定量的に推定する取り組みを行っており [4]，スマートフォンから得られる 128 次元の特徴量を用いて，ユーザのストレスの高低を平均約 70% の精度で推定できることを確認している．しかしながら，スマートフォンで得られるログとユーザのストレス状態の関係は複雑であり，推定精度の向上のためにはより高次の情報を考慮することが不可欠であると考えられる．これまでの先行研究 [5]，[6] ではストレスの推定のため，スマートフォンとウェアラブルセンサを組み合わせたり，複数のバイオマーカーを組み合わせたりすることで入力データ数を増やすアプローチが主流であった．しかしながら，複数センサの装着やバイオマーカーの取得は利用者の利便性を低下する原因となる．

この課題に対し，本研究ではセンシングデータの追加ではなく，ストレスと関連する因子に着目することでストレスの推定性能を向上させることを目的とし，ユーザの知的活動能力を示す注意機能を用いたストレス推定手法を提案する．様々な環境の変化によって生じるストレス反応は必ずしも注意機能を低下させるものではなく，適切なストレスは注意機能を高めることが知られており [7]，ストレスを受けている状態であっても注意機能が高い場合と低い場合ではストレスの表れ方が異なると考えられる．同様に，ストレスが高い場合と低い場合のそれぞれにおいて，注意機能の発揮のされ方が異なると考えられる．以上のようにストレスと注意機能には相互作用があり，両者の状態を考慮することがストレス推定の精度を高めることにおいて重要であると考え，本研究ではこの相互作用を考慮してストレスと注意機能を推定する手法を提案する．

提案手法では，スマートフォンで取得可能なモーションセンサ（加速度，角速度など），環境センサ（照度，位置情報など）の値，および利用ログ（画面点灯，アプリ起動履歴など）を用いて，ユーザのスマートフォンの持ち運びや利用に関する特徴量を生成する．この特徴量を用い，先行研究 [4] で用いた自律神経機能のバランスを反映する指標（LF/HF）を基にストレスの推定モデルを構築し，さらに認知的コントロール機能を計る Go-NoGo 課題 [8] により得られる注意機能の定量的なスコアに基づいて注意機能の推定モデルを構築する．これらのストレス推定モデルと注意機能推定モデルに加え，両者の相互作用を考慮し推定を行うモデルを構築することで，スマートフォンのログ，およびストレスと注意機能の相互作用に基づいたストレスの推定を行う．

¹ 株式会社 NTT ドコモ

² 東京大学

³ 東京大学 大学院工学系研究科 人工物工学研究センター

⁴ 慶應義塾大学 文学部 心理学専攻

⁵ 慶應義塾大学 医学部 精神・神経科学教室

以上の提案手法について、実際に 32 名の研究参加者から収集したのべ 500 日以上のデータセットを用いて評価を行った結果、ストレスと注意機能のそれぞれにおいて、両者を独立に推定する場合と比較して、相互作用を考慮して推定を行うことで約 5% の精度向上を確認し、ストレスの推定において注意機能を考慮することが効果的であることが明らかになった。

本研究の貢献は下記の通りである。

- ストレスと注意機能の相互作用に着目し、ストレスを推定するための因子として人の注意機能を導入し効果の検証を行った。
- ストレスと注意機能の 2 変数の相互作用を考慮し推定する方法として、異なる教師データを用いてモデルを作り分け、推定結果を重みに基づいて統合する方式を提案した。
- 注意機能の推定結果をストレスの推定にフィードバックすることで、ストレスの推定において約 5% の推定精度向上を確認した。

2. 関連研究と本研究の位置づけ

2.1 スマートフォンを用いた人の内面状態推定

スマートフォンには慣性センサ、GPS など複数のセンサが搭載されておりユーザの行動を取得可能であるため、スマートフォンを用いてユーザの内面状態を推定する研究が近年数多く行われている。例えば、ユーザの気分 [9]、社会的不安 [10]、迷い [11]、暇かどうか [12] など、外面からは推し量ることが難しい情報も、スマートフォンの持ち歩き方や利用傾向を用いて推定することが可能であることが示されている。

また、近年では医療分野に関連したユーザ状態の推定の研究も行われており、血糖値 [13]、双極性障害 [14]、うつ病の重症度 [15] などをスマートフォンの利用を通じて推定できる可能性が示されている。さらに、スマートフォンを用いたストレス推定に関する研究も報告されている [5]、[16]。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。しかしながら、先行研究ではいずれもユーザのアンケート回答結果をユーザのストレスの診断結果と位置づけており、ユーザの主観的なストレス状態を推定するに留まっている。したがって、ユーザが自覚していないストレスを推定することができず、ストレスに起因する精神疾患の予防において不十分である可能性が残る。

2.2 本研究の位置づけ

以上の先行研究に対し、本研究では定量的にストレスを計測する方式の確立を目的とし、スマートフォンから取得可能な情報のみを用いてユーザのストレスを推定する手法の検討をこれまでに行ってきた [4]。スマートフォンはユー

ザが毎日利用するものであることから、本研究ではスマートフォンを用いてユーザのストレスの経過を日単位で長期的に観察し、不調に至る前に事前にユーザに気づかせ、セルフケアの意識を高めることを目的としている。この目的のために、ユーザが自身のストレスを直感的に理解できることが重要であると考え、その日のストレスがユーザが平均的に受けているストレスより高いか低いかを推定する手法を検討する。本稿では著者らの先行研究 [4] を発展させ、ストレス推定精度のさらなる向上を目的として、ストレス推定に影響する因子として注意機能に着目し、機械学習的なアプローチを用いて両者の相互作用をモデル化する手法を提案する。ストレスに関連する人の内面状態を同時に推定しストレスを推定するアプローチは我々の知る限りこれまでになく、新しいアプローチである。

3. データセットの収集

スマートフォンログを用いてユーザのストレス、ならびに注意機能を推定する方式の検討のため、本研究では株式会社 NTT ドコモの勤務者を対象にデータセットの収集を実施した。データ収集には合計 39 人が参加し、合計のべ 1349 日のスマートフォンログデータを収集した。

データセットは (1) ストレス指標と注意機能指標の計測値、および (2) スマートフォンのログデータで構成されている。研究参加者はストレスの指標の計測のためにウェアラブル心電計 WHS-1*1 を胸部に装着し、注意機能の計測とログデータ収集のために開発した Android アプリを導入した。収集するログデータは位置情報やアプリケーション利用ログなどプライバシー情報を含むため、これらの情報は収集と同時に暗号化により保護される。また、研究参加者には ID を付与し、個人を特定できない形でデータ収集を行った。以上の本データ収集は、東京大学大学院医学系研究科・医学部倫理委員会の承認のもと実施した。

ストレス計測のための心電計の装着、ならびに注意機能計測アプリの実施は日常生活に支障がない範囲で任意で実施してもらい、それぞれのべ 638 日分の心電計データ、およびのべ 779 日分の注意機能計測結果を収集した。本研究では以上のうち、スマートフォンログ、心電計データ、注意機能計測結果の全てが揃っている日のデータを分析対象とし、合計のべ 554 日分のログデータを用いて 4 章のモデルの構築、および 5 章の評価を行う。

以降ではストレス・注意機能の指標の収集方法、ならびに収集したスマートフォンログの詳細を述べる。

3.1 ストレスの指標

本研究では、胸部心電計 WHS-1 を用いて計測した心電波形の低周波成分 (LF)、高周波成分 (HF) の比率を表す

*1 http://uniontool.co.jp/product/sensor/index_020101.html



図 1 Android アプリ上に実装した Go/NoGo 課題

指標 LF/HF[17] をストレスの指標として用いる。LF/HF は交感神経と副交感神経のバランスを表しており、値が高いほど交感神経が優位である、すなわち緊張状態にあることを表している。このように LF/HF は人の自律神経の状態を表す一方で、激しい運動や喫煙などの行動も LF/HF に影響を与えることが知られている。本研究では、デスクワークを中心とした社員を対象としてデータ収集を行っていること、研究参加者のうち喫煙者の割合が小さいこと、および喫煙時間に対し LF/HF の計算に用いる時間が十分長いことから、LF/HF をストレスの指標として扱う。

LF/HF の計算のために、心拍の各拍動の間隔 (RRI) のログ系列を 5 分ごとに分割したうえで、高速フーリエ変換を適用し周波数解析を実施しパワースペクトルに変換する。パワースペクトルのうち、0.04Hz から 0.15Hz までの領域を積分した値を LF、0.15Hz から 0.40Hz までの領域を積分した値を HF として計算し、HF に対する LF の比の値を LF/HF として計算する [18]。このようにして得られた 5 分ごとの LF/HF の値は、交感神経と副交感神経のバランスを表しており、時間の経過により細かく変化する。本研究では、2.2 節で述べた通り、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いるため、多くの研究参加者がデスクワークに従事している時間帯である 13 時から 16 時までの LF/HF の値の平均をユーザのその一日の LF/HF の代表値として扱う。

また、本研究では 2.2 節で述べた通り当日のストレスを高いか低いかの二値で推定するため、各ユーザの全日の LF/HF 代表値の平均値を基準値とし、当日の LF/HF の代表値が基準値以上の場合に 1 (高)、基準値未満の場合に 0 (低) のラベルをそれぞれ付与する。本方式でのべ 554 日のデータに対しラベル付けを行った結果、1 が 240 日、0 が 314 日となった。

3.2 注意機能の指標

本研究ではユーザの注意機能を計測する方法として、Android アプリ上に実装した Go/NoGo 課題 [8] を用いる。図 1 に実装した Go/NoGo 課題アプリケーションを示す。Go/NoGo 課題ではスマートフォンの画面に一定時間ごとにアルファベットが表示され、ユーザは所定の文字に対し

表 1 収集したスマートフォンログ (センサ計測値および利用ログ)

センサの計測値	
センサ種別	取得間隔
加速度センサ (3 軸)	1 秒
加速度センサ (重力成分除去) (3 軸)	1 秒
角速度センサ (3 軸)	1 秒
傾きセンサ (3 軸)	1 秒
気圧センサ	5 分毎
照度センサ	5 分毎
バッテリー残量	5 分毎
WiFi 接続先アクセスポイント (AP) 情報	5 分毎
周辺 WiFi アクセスポイント (AP) 情報	5 分毎
Google Activity Recognition	5 分毎
位置情報	5 分毎
ストレージ空き容量	毎日 23 時
カレンダー情報	毎日
端末利用ログ	
イベント種別	取得タイミング
充電器への接続	イベント発生時
イヤホンジャックの着脱	イベント発生時
画面のオン/オフ	イベント発生時
アプリの起動/終了	イベント発生時

できるだけ早く画面をタップ (Go 反応) し、その他の文字に対しては画面をタップしないこと (NoGo 反応) を求められる。アプリケーションでは約 1 分間で合計 72 回の文字 (予め定められた 8 種類の文字のうち 1 つ) が表示され、そのうち正しく反応できた Go 反応、および NoGo 反応の数の合計を 72 で割った値を本試験における注意機能のスコアとする。

ユーザはこの Go/NoGo 課題を一日あたり 3 回 (9:30-10:30, 12:00-13:00, 16:00-17:00 のそれぞれで 1 回ずつ) 試行し、3 回の試行で得られた注意機能スコアの平均値を当日の注意機能スコアとする。さらに、ストレスと同様に、各ユーザの全日の注意機能スコアの平均値を基準とし、当日の注意機能スコアが基準値以上の場合に 0 (高)、基準値未満の場合に 1 (低) のラベルを付与する。本方式により合計 554 日のデータについてラベルを付与した結果、1 が 224 日、0 が 330 日となった。

3.3 スマートフォンログ

前節で述べた Go/NoGo 課題のため実装した Android アプリケーションは、バックグラウンドでスマートフォンのログを収集する機能を有する。本研究で収集したスマートフォンログの一覧を表 1 に示す。ログの種類は二種類あり、スマートフォン内蔵のセンサ計測値、およびユーザがスマートフォンを利用したときに生じる端末利用ログである。センサの計測値は定められた間隔ごとに取得され、利用ログは当該イベント発生のたびに記録される。表 1 のセンサ値における加速度、角速度、傾き、気圧、照度、位置情報、バッテリー情報およびストレージ空き容量は、センサ

で取得した値をログに書き込んでおり、アクセスポイント情報はSSID名、Google Activity Recognition（以降GARと示す）は認識したActivity名、カレンダー情報は各曜日のいずれか、および平日か休日かの情報を記録する。また、利用ログでは充電、イヤホンについてはイベントの発生時間と接続・接続解除のいずれかを記録し、画面のオン/オフでは時間とオン・オフのいずれかを記録する。アプリの起動/終了については、イベントの発生時間と、起動/終了の種別、およびアプリ固有の識別子（パッケージ名）を記録する。

4. 提案手法

本章では、前章で収集したスマートフォンログを用いてストレス、および注意機能を推定する手法について述べる。提案手法の概要を図2に示す。提案方式では、ユーザの所持するスマートフォンログを用いて、まずユーザの行動の特徴を表す特徴量を生成する。特徴量の設計においては、人のストレスや注意機能の変化と関係のある身体の動きや外出傾向などの特徴を捉えるため、スマートフォンログから得られる複数のセンサ読み取り値をそれぞれ適した形式に変換し集計する。次に、生成した特徴量に対し、前章で計測したストレスと注意機能の指標を教師データとした機械学習によりストレスと注意機能の推定モデルを構築する。さらに、ストレスと注意機能の相互作用を考慮するために、注意機能状態に応じたストレス推定モデルと、ストレス状態に応じた注意機能推定モデルを用い、それぞれ注意機能の推定結果に基づいてストレスを再推定、およびストレスの推定結果に基づいて注意機能を再推定することで最終的な推定結果を出力する。

4.1 行動特徴量の生成

人が受けているストレスや注意機能の高低は人の行動に影響を与える。本研究では、人の行動の変化を表1で収集しているスマートフォンログを用いて捉えることを目指しており、以下の4つの観点と想定のもとでスマートフォンのログを特徴量化する。

(1) 人やスマートフォンの動きに関する特徴 – 人の心理状態が動きに反映される。(2) 周辺環境の特徴 – 周辺環境の変化が人に影響を与え、また、環境の変化で行動の変化を捉えられる。(3) 人の移動に関する特徴 – 人の心理状態が人の移動方法や訪れる場所に影響を与え、また、訪れる場所により人の心理状態が変化する。(4) スマートフォンの使い方の特徴 – 人の心理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。これら4観点に基づく特徴量化において、本研究では特徴量生成を1時間、6時間、24時間と時間的に異なるスケールのウィンドウを用いて行うことで、異なる粒度で人の行動の変化を捉えることを目指す。以上の方針で生成する特徴量の一覧を表2に示す。以

降では、(1)から(4)の特徴量の詳細を述べる。

4.1.1 人やスマートフォンの動きに関する特徴量

スマートフォン内蔵の動きに関するセンサ（加速度、角速度、傾き）を用いてスマートフォンを保持する人の動きや、スマートフォンそのものの動きを特徴量化する。動きの特徴を表す統計量として、計測値の平均値、標準偏差、最大値、最小値、最大値と最小値の差、各センサのベクトルの大きさ、各センサごとの各軸のペアについてのPearsonの相関係数をそれぞれ計算する。

4.1.2 周辺環境の特徴量

スマートフォン内蔵の気圧計や照度計の計測値を用いて平均値、標準偏差、最大値、最小値、最大値と最小値の差を計算する。気圧計や照度計の統計値を比較することで、当日のおおまかな天候、日照の度合いや、外出していたか、部屋の照明をいつまで利用していたかといった情報を得ることができる。

さらに、WiFiの接続情報、および周辺のWiFiアクセスポイント情報を利用する。特徴量抽出の対象時間のうち、WiFiが有効になっていた割合、WiFiに接続できていた割合、接続したWiFi APの数と、最も接続していたWiFi APへの接続回数を計算し、安定して通信を行える環境にどの程度滞在していたかを数値化する。また、収集した周辺WiFi情報について、対象時間のうち観測したWiFi AP名を重複を含む場合と重複を除く場合の2通りで集計することで、WiFi APが多数存在する環境（市街地など）を訪れたかどうかと、同じ場所に滞在していたかどうかの情報を得ることができる。

加えて、カレンダー情報を照合し、当日が何曜日か、および平日か休日かどうかを0または1の値で表現する。

4.1.3 人の移動に関する特徴量

Googleが提供するActivity Recognition APIを用いて、スマートフォンを保持する人の状態を推定することが可能である。推定される状態は、vehicle（車や電車）、bicycle（自転車）、on foot（歩行および走行）、running（走行）、walking（歩行）、still（静止）、tilting（傾きの変化）、unknown（その他）のいずれかである。特徴量抽出対象の時間のうち、これらの各クラスが発生した頻度を期間合計で1になるように正規化した値を特徴量として用いる。さらに、本方式ではこれらのクラスに加え、ride（vehicle + bicycle）、move（on foot + running + walking）という行動を定義し加える。

また、GPSやWiFiで測位した位置情報を2つの異なる方式で集計し特徴量化する。まず1つめの方式として、緯度、経度および高度の情報をを用い、最大値、最小値、最大値と最小値の差を計算する。この際、相対的な座標移動を考慮するため、各日の初回に計測した座標を基準とした値を用いる。加えて、緯度と経度の各計測タイミングにおける変化から移動距離を計算しその合計を得る。

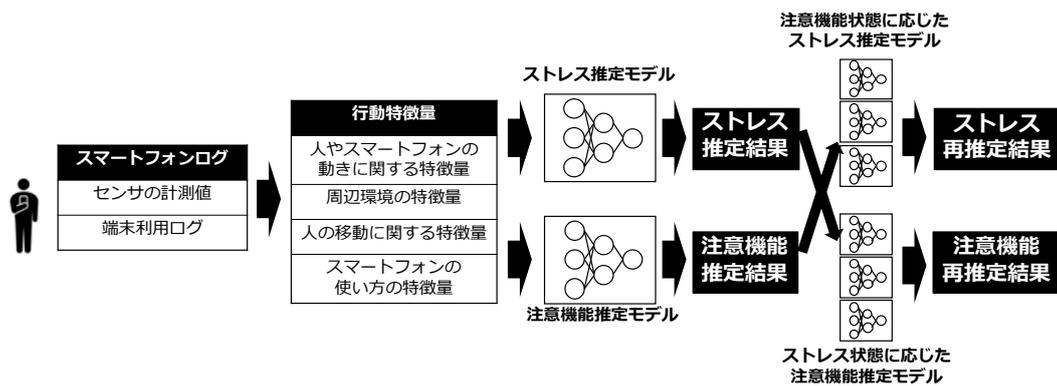


図 2 提案手法の概要

表 2 スマートフォンログから生成する行動特徴量

センサ・ログ種別	生成する特徴量	次元数	集計単位
加速度, 角速度, 傾き	平均, 標準偏差, 最大値, 最小値, 最大最小差, ベクトルの大きさ, 各軸の相関係数	2204	1 時間, 6 時間, 24 時間
気圧, 照度, バッテリ	平均, 標準偏差, 最大値, 最小値, 最大最小差	435	
接続先 AP, 周辺 AP	WiFi ON 率, WiFi 接続率, 接続 AP 数, 最接続 AP 接続回数, 周辺 WiFi 数	145	
GAR	vehicle, bicycle, on foot, running, walking, still, tilting, unknown, ride, move の各割合	290	24 時間
位置情報	緯度, 経度, 高度の最大値, 最小値, 最大最小差 (各日の初回計測値を基準とする), 緯度経度差から求めた移動距離合計	290	
	回転断面半径, 自宅からの最大距離, トリップ (合計, 最大値, 標準偏差), 滞留点数 (通常, 重要, 新規), 滞留点類似度 (前日比, 7 日前比)	12	
ストレージ空き容量	残容量の値	1	1 時間, 6 時間, 24 時間
カレンダー	各曜日かどうか, 平日かどうか	9	
充電器, イヤホン接続	接続回数, 接続されていた割合	116	
画面オン/オフ	画面オン回数, オフ回数, オンオフ回数合計, 点灯時間の割合 オン時間 (平均, 最大, 最小, 標準偏差)	232	24 時間
	1 日の合計利用時間, 利用時間の平均	2	
	1 時間毎 ON 頻度の類似度, 平均 ON 時間の差, 合計 ON 回数の差, 合計 ON 時間の差 (ユーザ平均比, 前日比, 7 日前比)	12	
アプリ	各カテゴリの 1 時間毎 ON 頻度の類似度, 平均 ON 時間の差, 合計 ON 回数の差, 合計 ON 時間の差 (ユーザ平均比, 前日比, 7 日前比)	384	1 時間, 6 時間, 24 時間
	各カテゴリおよび全カテゴリの起動時間 (平均, 標準偏差), 起動回数, コミュニケーション回数, 時間 (通話, SMS) セッション数, アプリ切替回数平均, セッション内アプリ利用数の平均, 最大, 最小	3305	

さらに 2 つめの方式として、位置情報よりユーザの滞留点を集計し特徴量化を行う。滞留点はユーザが滞在した地点を集めたものであり、一日の緯度経度の座標系列をクラスタリングし、重要な点の集まりをまとめることで計算する。滞留点のクラスタリングには時空間 mean shift clustering アルゴリズム [19] を用いる。以上の処理により得られた滞留点の系列を一日の移動経路とみなした上で、先行研究 [20] を参考に、回転断面半径、自宅からの最大距離、各滞留点間の移動 (トリップ) の合計距離、トリップの最大値、標準偏差を求め、移動の規則性や軌跡の形状を考慮する。さらに、滞留した点の合計数と、頻繁に訪れる重要な滞留点、および新しく訪れた滞留点の数をそれぞれ求め、ユーザが訪れた場所の多様性や新規性を評価する。最後に、当日の滞留点と、前日および 7 日前の滞留点の系列を比較し、移動経路がどの程度類似しているか、すなわ

ち行動の規則性を評価する。移動経路の類似性の比較には Jaccard Index および Dice Index を用い、一時間単位で計算した各指標の平均値を一日の類似度の指標とする。

4.1.4 スマートフォンの使い方の特徴量

スマートフォンが使われる目的や使われ方を解釈するため、以下の特徴量を設計する。バッテリーの残量について、各時間帯の平均値、標準偏差、最大値、最小値、最大値と最小値の差を計算し、過度にバッテリーを消費してでも使いたいかどうかや、バッテリーが原因でスマートフォンを使わなかったかどうかの情報が得られる。また、充電器への接続回数や、接続されていた時間割合を集計することで、ユーザにとってスマートフォンを充電しておく必要が生じたかどうかや、通常充電を行う時間帯の傾向を知ることができる。

さらに、一日単位でのストレージ容量の変化を比較す

ることで、多数の写真の撮影や音楽のダウンロードなど、ユーザの趣味、娯楽に関する情報が得られることを期待している。同様に、イヤホンジャックへの接続回数や接続時間の割合を知ることで、音楽を聞くという娯楽に関する活動時間を推し量ることができる。

画面の点灯はスマートフォンが利用されていることを表しており、本方式では頻度的な側面と時間的な側面の両方で特徴量化を行う。頻度を表す特徴量として画面オンの回数、画面オフの回数、および画面オンとオフの回数を計算し、時間を表す特徴量として、画面オンが継続する時間の平均値、標準偏差、最大値および最小値、一日の合計利用時間、ならびに集計時間に占める画面点灯時間の割合を計算する。加えて、一日を通じた画面点灯時間の合計および平均を計算する。さらに、一日を通じてどの時間帯に画面を点灯させたかを計るため、1時間ごとの画面オン回数を24時間合計のオン回数で割ることで、一日における各時間帯における画面オンの頻度を求める。さらに、当日における各時間帯の画面オンの頻度分布をユーザの平均的な分布、前日の分布、ならびに7日前の頻度分布と比較することで、画面点灯パターンの類似度を数値化する。前述の類似度の比較においてはバタチャリヤ距離 [21] を用いる。

アプリケーションの利用はスマートフォンを利用する主目的である。本手法では、アプリごとに定められたカテゴリをユーザがアプリを利用する目的とみなし、アプリごとに異なる名称（識別子）ではなく、カテゴリ単位でアプリログを集計する。本研究では32人の研究参加者が利用した上位32個のカテゴリ（OS標準の通話機能、コミュニケーション、エンターテインメント、ツールなど）を集計対象とし、各カテゴリの起動時間（平均、標準偏差）、および起動回数を集計する。加えて、全てのカテゴリの起動時間および起動回数を集計する。さらに、利用のパターンを特徴量化するために、各カテゴリの1時間毎の起動回数の頻度割合（一日における各カテゴリの起動頻度分布）、1時間ごとの各カテゴリの起動回数割合（各時間帯における全カテゴリの起動頻度分布）、を計算する。加えて、画面点灯パターンの類似度の計算と同様に、各カテゴリの1時間毎の起動頻度分布をユーザの平均的な分布、前日の分布、ならびに7日前の頻度分布と比較する。また、スマートフォンの主な機能である通信に着目し、通信のためのアプリ（通話およびショートメッセージ）に注目し特徴量を追加する。具体的には、これらのアプリの中でも電話の発信/着信、メッセージの送信/受信といった画面に遷移した回数と画面での経過時間を集計し、コミュニケーションの回数とコミュニケーション実施時間として集計する。

また、スマートフォンでは複数のアプリを切り替えながら使われることも多いと考え、時系列的に連続して起動されているアプリログを一つの単位（セッション）とみなし、ユーザが各セッションにおいてどの程度アプリを切り替え

て利用しているかといった情報を計算する。具体的には、ウィンドウ内のセッションの総数、各セッションにおけるアプリ切り替え回数の平均値、セッション内アプリの平均値、最大値、最小値を特徴量として計算する。なお、セッション単位での集計においてはアプリはカテゴリ単位ではなく名称単位で集計し、同一カテゴリで異なる名称のアプリを連続して利用した場合もアプリの切り替えとしてみなす。

4.2 ストレスと注意機能の推定

4.1節で述べた合計16種類のログデータより生成した合計7437次元の特徴量を説明変数とし、3.1節、3.2節で述べたストレスおよび注意機能のラベルを目的変数とした機械学習アルゴリズムによりストレスと注意機能の推定モデルを構築する。機械学習の手法は種々存在するが、本研究では近年数多くの研究で優れた性能を示しているXGBoost[22]を用いる。

機械学習アルゴリズムにおいては、入力次元数が多いと複雑な表現が可能になる一方で、計算時間の増大や過学習の原因となることが知られているため、用いる特徴量の選択が重要である。XGBoostでは訓練したモデルにおいて、各特徴量の推定における重要度（feature_importance）を計算可能であり、本方式ではこの重要度を用いて特徴量選択を行う。具体的には、訓練データにおける重要度全体の平均値を計算し、平均値を超える特徴量のみを抽出し、再度モデル構築を行う処理を繰り返すことで特徴量を選択し、高い精度が得られるモデルを構築可能な特徴量を探索する。以上で構築したストレスと注意機能のそれぞれの推定モデル（ M_{stress} , M_{alert} と示す）を用いることで、ストレスと注意機能のそれぞれを推定することが可能になる。本研究ではストレスが高い状態、注意機能が低下した状態をそれぞれ陽性と定義し、モデル M_{stress} , M_{alert} はそれぞれ、ストレスが高いクラスに所属する確率 p_1 、注意機能が低いクラスに所属する確率 q_1 を出力し、これらの値が0.5以上となった場合に陽性と判断する。

4.3 ストレスと注意機能の相互作用を考慮した再推定

4.2節で推定したモデル M_{stress} , M_{alert} では、ストレスと注意機能の相互作用については全く考慮していない。すなわち、ストレスと注意機能をそれぞれ独立に推定することができる。一方で、ストレスを受けた際に分泌が増えるコルチゾールが人の注意機能に作用することが分かっている [23]、ストレスの状態を踏まえて注意機能を推定すること、および注意機能の状態を踏まえてストレスを推定することが有効であると考えられる。しかしながら、ストレス、注意機能が相互に影響し合うメカニズムについて、生体を対象とした実験的研究に基づいて厳密にモデル化することは簡単ではない。

そこで、本研究では機械学習的なアプローチに基づいて相互作用を推定アルゴリズムに組み込む方式を提案する。提案方式では、ストレスが高い場合、ストレスが低い場合を想定した注意機能の推定モデル $M_{alert_Hs}, M_{alert_Ls}$ と、注意機能が高い場合、注意機能が低い場合を想定したストレスの推定モデル $M_{stress_Ha}, M_{stress_La}$ を教師データを使い分けることによって作り分ける。例えば M_{stress_Ha} の構築においては、注意機能が高いときのログデータのみを集計し、ストレスが高いか低いかの2値を推定する分類問題を学習させる。このようにシチュエーションに応じてモデルを作り分けることで、シチュエーションに適合した場合におけるストレスや注意機能の推定精度向上が期待できる。一方で、シチュエーションに沿わない場合には推定精度が低下する恐れがあるため、正確にシチュエーションを特定することが重要である。そこで、提案手法では前節で独立に推定したストレスおよび注意機能の陽性確率 p_1, q_1 をシチュエーションの判断に用い、 p_1, q_1 の値が中立、すなわちシチュエーションの判断が難しい場合にはシチュエーションを定めない汎用的な推定結果である p_1, q_1 を利用するアプローチを採用する。以上の方針のもとで、 p_1, q_1 の値に応じて各モデルの推定結果を適応的に組み合わせる方式として、次式 (1), (2) によりストレスと注意機能の陽性確率 p, q を再計算する。

$$p = w_1 * p_1 + w_2 * p_2 + w_3 * p_3 \quad (1)$$

$$q = w'_1 * q_1 + w'_2 * q_2 + w'_3 * q_3 \quad (2)$$

ただし、式中 p_2, p_3, q_2, q_3 はそれぞれモデル $M_{stress_Ha}, M_{stress_La}, M_{alert_Hs}, M_{alert_Ls}$ が出力する陽性確率とする。各式における w_1 から w'_3 はストレスと注意機能の再計算における、各推定モデルの出力した陽性確率の足し合わせの重みを表しており、以下の式で計算される。

$$w_1 = 2 - 2 \cdot q_1, w_2 = 0, w_3 = 2 \cdot q_1 - 1 \quad (\text{if } q_1 \geq 0.5) \quad (3)$$

$$w_1 = 2 \cdot q_1, w_2 = 1 - 2 \cdot q_1, w_3 = 0 \quad (\text{if } q_1 < 0.5) \quad (4)$$

$$w'_1 = 2 - 2 \cdot p_1, w'_2 = 2 \cdot p_1 - 1, w'_3 = 0 \quad (\text{if } p_1 \geq 0.5) \quad (5)$$

$$w'_1 = 2 \cdot p_1, w'_2 = 0, w'_3 = 1 - 2 \cdot p_1 \quad (\text{if } p_1 < 0.5) \quad (6)$$

各重みは p_1, q_1 を基準に場合分けを行った上で計算される。各指標が陽性と判断される区間 $[0.5, 1]$ 、および陰性と判断される区間 $[0, 0.5)$ の2通りの場合分けが行われ、各場合において計算される重み w_1 から w_3 、および w'_1 から w'_3 は各和が1となる式で求められる。

式 (3), (4) はそれぞれ注意機能が低い場合、高い場合のストレス再推定における重みを表している。式 (3) は注意機能が低い想定のため、注意機能が高い想定 of ストレス推定結果 p_2 に対する重み w_2 の値はゼロとし、再推定に用いないことを表している。また、 w_1, w_3 は q_1 の値によって変動する式となっており、 q_1 の値が1に近い、すなわち注

意機能が低い確率が高いほど、注意機能が低い想定 of ストレス推定結果 p_3 に係る重み w_3 が大きくなる。この際、 q_1 を2倍した値を計算に用いることで、区間 $[0.5, 1]$ における w_1, w_3 の値域をそれぞれ $[1, 0], [0, 1]$ とする。逆に、 q_1 の値が0.5に近く注意機能推定の確率が低いほど、注意機能と独立にストレスを推定した結果 p_1 の重み w_1 が大きくなる。同様に、(4) は注意機能が高い確率が高いほど、注意機能が高い想定 of ストレス推定結果 p_2 に係る重み w_2 を大きくする式である。

上記の式 (3), (4) と同様に、式 (5), (6) はストレスが高い想定、低い想定 of それぞれを表しており、注意機能と独立に推定したストレスの推定結果 p_1 の確率に応じて、注意機能の推定結果 q_1, q_2, q_3 を足し合わせる重み w'_1, w'_2, w'_3 を調整する式となっている。

以上の式により、一度独立に推定したストレスと注意機能の推定結果と確率を基に、注意機能とストレスの相互作用を考慮した再推定を行える。なお、式 (1), (2) により推定結果 p, q の値が区間 $[0, 1]$ を超える場合があるが、0.5以上を陽性、それ以外を陰性と判断する。

5. 性能評価

本章では、4章で述べたストレスと注意機能の推定方式について、3章のデータセットを用いて性能評価を行った結果を述べる。まず5.1節で性能評価の方法を述べ、5.2節で提案手法によるストレスと注意機能の推定性能を評価する。最後に、5.3節でストレスと注意機能の推定に寄与した行動特徴量について議論し、本研究のアプローチ方法の妥当性を検証する。

5.1 評価環境

性能評価にあたって、3章で述べた32人ののべ554日分のデータセットについて、まず異なるユーザ間の特徴量の正規化と欠損した特徴量の補完を実施した。特徴量の正規化のため、各ユーザ単位でデータセットを分割し、欠損値を除いた上で、特徴量ごとに平均が0、分散が1の分布に従うようにデータを変換する。その後、各特徴量における欠損値について、ユーザの平均値を用いた補完を行う。ある特徴量について、ユーザの特徴量が全て欠損している場合においては、その他の全ユーザの平均値を用いて補完を実施する。

以上の方法で正規化した特徴量と、ストレスと注意機能のラベルを合わせて、32人のユーザに対して一人抜き交差検証を適用する。一人抜き交差検証では、システムの評価対象とするユーザ一人を定め、その他のユーザのデータセットを用いてシステム（機械学習モデル）を訓練することで、未知のユーザに対するシステムの性能を評価する。したがって、本研究では32回のデータ分割を実施し32人それぞれが未知のユーザであると想定した評価を行い、各

表 3 ストレスと注意機能の推定結果

推定方式	推定対象	性能指標			
		accuracy	specificity	sensitivity	precision
(1) ベースライン	ストレス	50.9%	56.7%	43.3%	43.3%
	注意機能	52.0%	59.7%	40.6%	40.6%
(2) 独立推定	ストレス	75.3%	79.3%	70.0%	72.1%
	注意機能	79.1%	81.8%	75.0%	73.7%
(3) 提案手法	ストレス	80.7%	82.8%	77.9%	77.6%
	注意機能	80.9%	86.1%	73.2%	78.1%
(4) 上限値	ストレス	81.2%	83.4%	78.3%	78.3%
	注意機能	83.2%	87.2%	76.8%	80.4%

ユーザの推定結果を合わせて評価結果とする。評価結果を計る性能指標として、accuracy（陽性および陰性の正解率）、specificity（陰性の検出率）、sensitivity（陽性の検出率）、precision（陽性検出の成功率）をそれぞれ用いる。

また、機械学習モデルの訓練において陽性と陰性のデータ数に偏りが生じる場合があり、偏った教師データで学習することは過学習の原因となるため、SMOTE(synthetic-minority-over-sampling-technique)[24]を用いてオーバーサンプリングを行い、教師データにおける陽性と陰性の割合を均一化した。

以降の性能評価では、次の4通りの推定方式を比較し、提案手法の有効性について議論する。(1) ベースライン：スマートフォンログを用いず、評価データのストレス・注意機能の高低それぞれの確率に従い、ランダムにストレスと注意機能を推定する場合の性能を示す。(2) 独立推定：4.2節で述べた、ストレスと注意機能をそれぞれに対して独立に推定モデルを構築し推定する方式。(3) 提案手法：4.3節で述べた、ストレスと注意機能の推定結果に基づいて、合計6種類のモデルの推定結果を用いストレスと注意機能を再計算する方式。(4) 上限値：(3)の提案手法において、再計算の前提となるストレスと注意機能の推定結果 (p_1, q_1) について、正しい推定結果（真値）を用いる方式であり、提案手法において、注意機能の状態を正しく推定できた前提における提案手法によるストレスの推定結果、およびストレスの状態を正しく推定できた前提における提案手法による注意機能の推定結果を示しており、相互作用を考慮することにより得られる推定性能の上限値を示す。

5.2 評価結果

5.2.1 全体の平均性能

表3に各手法で推定したストレスと注意機能の推定性能を示す。表中(1)と(2)の比較により、ストレス推定の各性能指標がランダムな推定と比較して20%以上向上しており、筆者らの先行研究と同様に、スマートフォンログを用いてストレス推定モデルを構築する手法の有効性が確認できる。また、注意機能の推定においてもスマートフォンログを用いた推定の有効性を示している。

また、表中(2)と(3)の比較により、注意機能の推定結

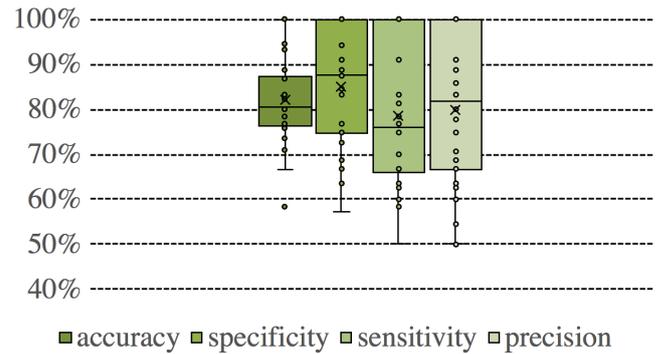


図 3 ユーザ単位での性能評価結果

果を踏まえてストレスの推定結果を再計算する提案手法によりストレスの推定精度 (accuracy) が 5.4% 向上したことから、ストレスの推定において注意機能の状態を考慮することが有効であることが確認できる。また、sensitivity, specificity の両指標が向上していることから、注意機能の考慮により、ストレスが高い場合と低い場合のいずれにおいても推定精度を向上させられたことが示されており、提案手法により分類問題として全体的に精度の改善を達成できたことが分かる。さらに、ストレスの推定における precision は 77.6% であり sensitivity と同等の精度であることから、陽性の検出における適合率、再現率のトレードオフが両立しており、ストレスの推定モデルとして学習が適切に行われていることが確認できる。この結果は一人抜き交差検証で得られたものであり、評価したいユーザのデータを用いずとも、同様のワークスタイルの他のユーザのデータを用いて推定モデルを構築することが可能であることが示唆されている。

さらに、表中(3)と(4)の比較により、注意機能の真値が明らかになった上でのストレス推定で得られる性能上限値に対し、提案手法では各指標において高々 0.6% の推定性能低下に留まっており、ほぼ理想的な性能が得られたことが分かる。一方で、上限値においても未だストレスの推定性能は 81.2% であり、スマートフォンのみで得られる情報から複数のユーザのストレスの高低を確実に推定することは容易ではないことも示されている。この原因として、精神的ストレスを受けた際の行動（特徴量）の変化がユーザによって異なった結果、現在の特徴量ではストレスの状態を完全に推定することが難しかったものと考えられる。

この問題に対し、ストレスを受けた際の一般的な特徴量の変化に加え、個々のユーザにおいて独特な行動の変化を考慮することが有効であると考えており、今後、ユーザの生活パターンを学習し、個人に合わせて推定モデルをチューニングする方式を検討している。

5.2.2 ユーザごとの平均性能

さらに、ユーザごとの推定性能の平均値の箱ひげ図を図

3に示す。ユーザによっては実験データの収集期間が短く、ストレスの真値が陽性または陰性のみ存在し、specificity, または sensitivity の値を評価できないユーザが存在する。したがって、本評価ではストレスの真値が陽性または陰性のみ存在するユーザ2名を除き、30名のユーザの各性能評価指標を示す。得られた30名ののべ551日分のデータセットにおける各ユーザのデータ数の平均は18.4日、標準偏差は8.6日である。

図3中の丸印は各ユーザの指標の値を示しており、バツ印は全ユーザの平均値を示している。この結果より、いずれの指標においても最高性能は100%を示しており、ユーザによってはスマートフォンの利用ログのみを用いてストレスの高低の推定を正確に行えることが示されている。また、各性能指標 {accuracy, specificity, sensitivity, precision} の最低値はそれぞれ {58.3%, 57.1%, 50.0%, 50.0%} であり、提案手法により少なくとも表3におけるベースライン以上の推定パフォーマンスを得られており、30人の異なるユーザにおいても提案手法により一貫してストレスを推定可能であることが示唆されている。

5.3 各指標の推定に寄与した特徴量の分析

5.2節の評価により、ストレスと注意機能の相互作用を機械学習的なアプローチに基づいて推定アルゴリズムに組み込む提案手法の有効性を示した。本節では、機械学習的な観点に基づいて上記の結果の妥当性を検証する。提案手法で構築した合計6種類の推定モデルにおいて、特徴量選択の結果得られた推定における重要度の高い特徴量の上位10件を表4に示す。なお、実験で用いたスマートフォンの軸は、画面に向かって右方向がX軸、上方法がY軸、手前方向がZ軸を表す。

表4の結果より、6種類のモデルはそれぞれ異なる特徴量を参照してストレス・注意機能の推定を行っていることが分かる。表中では、ストレスの推定や注意機能の推定において、一部共通する特徴量を用いている場合もあるが、注意機能が高い場合と注意機能が低い場合のストレス推定、ならびにストレスが高い場合とストレスが低い場合の注意機能推定のいずれにおいても推定に寄与する特徴量は共通しておらず、提案手法によって作り分けたモデルが異なる特徴量を参照して推定を行っていることがわかる。また、傾向として、注意機能が高いときのストレス推定、およびストレスが高いときの注意機能推定(表中央列)においてアプリ利用に関する特徴量が利用されている一方で、注意機能が低いときのストレス推定、およびストレスが低いときの注意機能推定(表右列)ではアプリ利用の特徴量が選択されておらず、端末の動きや傾きに関する特徴量がほとんどを占めている。このことから、注意機能やストレスが高いときはスマートフォンの使い方にユーザの心理状態が反映される一方で、注意機能が低下しているときやストレス

が低いときはスマートフォンの使い方ではなく身体やスマートフォンの動きでユーザの心理状態を推定可能であることが分かる。さらに、注意機能が低い場合のストレス推定(表右上)では加速度や傾きの最大値や偏差など、動きの激しさやばらつき度合いが推定に寄与することに対し、ストレスが低い場合の注意機能推定(表右下)では、加速度や角速度の相関係数が上位を占めており、動きの規則性が推定において重要であることを示している。

以上の結果より、ストレスが高いとき、低いとき、ならびに注意機能が高いとき、低いときのそれぞれにおいて、各モデルが目にする特徴量が異なっており、それぞれの状態に特化したモデルになっていることが分かる。その結果、独立に推定するモデルのみを用いる場合と比較して、ストレスまたは注意機能の状態に応じて、注目すべき特徴量を変えながらモデルの推定結果を適応的に組み合わせる提案手法の推定精度が向上したものと考えられる。

したがって、ストレスと注意機能をスマートフォンログを用いて推定する本課題において、ストレスと注意機能の推定モデルの作り分け、ならびに各モデルの推定結果の組み合わせで両指標を再度推定する提案手法が、ストレスと注意機能の相互作用を考慮するアプローチとして妥当であったといえる。

6. まとめと今後の課題

本研究では、スマートフォンを用いたユーザのストレス状態推定においてユーザの注意機能を重要因子であると位置づけ、ストレスと注意機能の相互作用をモデル化し、注意機能の推定結果に基づいてストレスを推定する手法を提案した。実際に、32人の会社員より集めたのべ554日分のスマートフォンログを用いて提案手法の評価を行った結果、平均80.7%の精度でユーザのストレスの高低を推定可能であることが明らかになった。さらに、注意機能を考慮しないストレスの推定精度に対し、提案手法による推定精度が5.4%向上することを確認し、ストレスの推定において注意機能が有効な因子であることを確認した。

今後の課題として、提案手法の適用範囲を拡大し、多様なユーザにおいて一貫してストレスを推定可能であることを示すことが重要であると考えており、会社員以外に多種多様なユーザにおける訓練データを収集したモデルの構築が必要であると考えている。また、ユーザごとに異なる生活パターンを考慮するため、蓄積した個人のデータに基づいて推定モデルをチューニングする方式を検討している。

謝辞 本研究の遂行にあたり、データセットの収集にご協力くださった方々にお礼申し上げます。

参考文献

- [1] 厚生労働省: 厚生労働省疾患調査, <https://www.mhlw.go.jp/toukei/saikin/hw/kanja/14/d1/toukei.pdf>.

表 4 各モデルが選択した重要な特徴量 (括弧内は時間帯を示す)

ストレス推定モデル		
注意機能と独立 (M_{stress})	注意機能が高い場合 (M_{stress_Ha})	注意機能が低い場合 (M_{stress_La})
通話アプリ起動回数の増減 (対ユーザ平均) (終日) 加速度 X・Z 軸相関係数 (7) 高度最大値 (12) 端末傾き標準偏差 (16) 端末傾き標準偏差 (15) 端末傾き平均 (18-24) 角速度 Y 平均 (6-12) 加速度 X・Y 軸相関係数 (3) 角速度 X・Z 軸相関係数 (16) 加速度 X・Z 軸相関係数 (19)	角速度 Y 平均 (6-12) 角速度 Y 最小値 (終日) 加速度 Y 平均 (10) 高度最小値 (12-18) 加速度 Y 標準偏差 (12-18) 通話アプリ起動回数の増減 (対ユーザ平均) (終日) 角速度 X・Y 軸相関係数 (8) 加速度 Y 平均 (17) 加速度 X・Z 軸相関係数 (19) 加速度ベクトルの大きさ (16)	加速度 X 最大値 (20) 加速度 Z 標準偏差 (0) 加速度ベクトルの大きさ (22) 端末傾き標準偏差 (5) 端末傾き最大値 (終日) 加速度 Y・Z 軸相関係数 (21) 角速度 X・Z 軸相関係数 (15) 端末傾き最小値 (終日) 端末傾き最大値 (15) 緯度最小値 (18-24)
注意機能推定モデル		
ストレスと独立 (M_{alert})	ストレスが高い場合 (M_{alert_Hs})	ストレスが低い場合 (M_{alert_Ls})
加速度 X・Y 軸相関係数 (1) 加速度 Y・Z 軸相関係数 (6-12) 角速度 X 最大値 (9) 加速度 X 平均 (1) 気圧の標準偏差 (2) 加速度 Y 標準偏差 (8) 角速度 Z 最小値 (12-18) 加速度 Y 最小値 (5) 加速度 Y 平均 (18-24) SOCIAL ¹ アプリ起動回数の増減 (対ユーザ平均) (終日)	SOCIAL ¹ アプリ起動回数の増減 (対ユーザ平均) (終日) 加速度ベクトルの大きさ (23) 端末傾き最小値 (22) 加速度 Y・Z 軸相関係数 (7) 端末傾き最大値 (6-12) 期間内の MARKET ² カテゴリの利用率 (13) 角速度 X・Z 軸相関係数 (0-6) 加速度 Z 最小値 (13) 加速度 X 最大最小差 (10) 角速度 Y・Z 軸相関係数 (6)	加速度 X・Y 軸相関係数 (1) 加速度 Y・Z 軸相関係数 (21) 加速度 Y・Z 軸相関係数 (終日) 端末傾き最小値 (6) 加速度 Y・Z 軸相関係数 (6-12) 端末傾き最大値 (18-24) 端末傾き標準偏差 (15) 角速度 Y・Z 軸相関係数 (1) 加速度最小値 (18-24) 角速度 Y・Z 軸相関係数 (21)

¹ Instagram, Facebook, LinkedIn, Tumblr など

² d マーケット, GALAXY Apps, Disney マーケット

- [2] 夏目 誠: 勤労者のストレス評価法 (第 2 報): ストレスドック受検者の 1 年間における体験ストレス点数の合計点とストレス状態や精神障害との関連から, 産業衛生学雑誌, Vol. 42, No. 4, pp. 107–118 (2000).
- [3] Ebert, D. D., Cuijpers, P., Muñoz, R. F. and Baumeister, H.: Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research, *Frontiers in psychiatry*, Vol. 8, p. 116 (2017).
- [4] Yamamoto, N., Ochiai, K., Inagaki, A., Fukazawa, Y., Kimoto, M., Kiriu, K., Kaminishi, K., Ota, J., Okimura, T., Terasawa, Y. and Maeda, T.: Physiological Stress Level Estimation Based on Smartphone Logs, *Proceedings of the 11th International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, IEEE (2018).
- [5] 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).
- [6] Aigrain, J., Spodenkiewicz, M., Dubuisson, S., Detyniecki, M., Cohen, D. and Chetouani, M.: Multimodal stress detection from multiple assessments, *IEEE Transactions on Affective Computing* (2016).
- [7] Sandi, C.: Stress and cognition, *Wiley Interdisciplinary Reviews: Cognitive Science*, Vol. 4, No. 3, pp. 245–261 (2013).
- [8] Hirose, S., Chikazoe, J., Watanabe, T., Jimura, K., Kunimatsu, A., Abe, O., Ohtomo, K., Miyashita, Y. and Konishi, S.: Efficiency of go/no-go task performance implemented in the left hemisphere, *Journal of Neuroscience*, Vol. 32, No. 26, pp. 9059–9065 (2012).
- [9] LiKamWa, R., Liu, Y., Lane, N. D. and Zhong, L.: Moodscope: Building a mood sensor from smartphone usage patterns, *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, ACM, pp. 389–402 (2013).
- [10] Fukazawa, Y., Ito, T., Okimura, T., Yamashita, Y., Maeda, T. and Ota, J.: Predicting anxiety state using smartphone-based passive sensing, *Journal of biomedical informatics*, p. 103151 (2019).
- [11] Narimoto, R., Kajita, S., Yamaguchi, H. and Higashino, T.: Wayfinding Behavior Detection by Smartphone, *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, IEEE, pp. 488–495 (2018).
- [12] Pielot, M., Dingler, T., Pedro, J. S. and Oliver, N.: When attention is not scarce-detecting boredom from mobile phone usage, *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, ACM, pp. 825–836 (2015).
- [13] Gu, W., Zhou, Y., Zhou, Z., Liu, X., Zou, H., Zhang, P., Spanos, C. J. and Zhang, L.: Sugarmate: Non-intrusive blood glucose monitoring with smartphones, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 1, No. 3, p. 54 (2017).
- [14] Grünerbl, A., Muaremi, A., Osmani, V., Bahle, G., Oehler, S., Tröster, G., Mayora, O., Haring, C. and Lukowicz, P.: Smartphone-based recognition of states and state changes in bipolar disorder patients, *IEEE Journal of Biomedical and Health Informatics*, Vol. 19, No. 1, pp. 140–148 (2015).
- [15] Canzian, L. and Musolesi, M.: Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis, *Proceedings of the 2015 ACM international joint conference on pervasive*

- and ubiquitous computing, ACM, pp. 1293–1304 (2015).
- [16] Bogomolov, A., Lepri, B., Ferron, M., Pianesi, F. and Pentland, A. S.: Daily stress recognition from mobile phone data, weather conditions and individual traits, *Proceedings of the 22nd ACM international conference on Multimedia*, ACM, pp. 477–486 (2014).
- [17] 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).
- [18] Rawenwaaij-Arts, C., Kallee, L., Hopman, J. et al.: Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability. Standards of measurement, physiologic interpretation, and clinical use. *Circulation* 1996; 93: 1043-1065., *Intern. Med*, Vol. 118, pp. 436–447 (1993).
- [19] Nishida, K., Toda, H., Kurashima, T. and Suhara, Y.: Probabilistic identification of visited point-of-interest for personalized automatic check-in, *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, pp. 631–642 (2014).
- [20] Wang, R., Aung, M. S., Abdullah, S., Brian, R., Campbell, A. T., Choudhury, T., Hauser, M., Kane, J., Merrill, M., Scherer, E. A. et al.: CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia, *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, pp. 886–897 (2016).
- [21] Cha, S.-H.: Comprehensive survey on distance/similarity measures between probability density functions, *City*, Vol. 1, No. 2, p. 1 (2007).
- [22] Chen, T. and Guestrin, C.: Xgboost: A scalable tree boosting system, *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, ACM, pp. 785–794 (2016).
- [23] Echouffo-Tcheugui, J. B., Conner, S. C., Himali, J. J., Maillard, P., DeCarli, C. S., Beiser, A. S., Vasani, R. S. and Seshadri, S.: Circulating cortisol and cognitive and structural brain measures: The Framingham Heart Study, *Neurology*, Vol. 91, No. 21, pp. e1961–e1970 (2018).
- [24] Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P.: SMOTE: synthetic minority over-sampling technique, *Journal of artificial intelligence research*, Vol. 16, pp. 321–357 (2002).