

3種の活動量計の比較による活動量認証モデルの評価

鈴木 宏哉^{1,a)} 山口 利恵^{1,b)}

受付日 2017年6月12日, 採録日 2017年12月8日

概要: 近年, オンラインサービスの増加に加え, モバイルデバイスの普及にともない, 個人認証技術の重要性が増している. 多要素認証技術の導入などによる安全性の向上が推奨されているが, 一方で認証要素の増加は利便性の低下という課題も含んでいる. 利便性を考慮した認証手法として, ユーザ自身による明示的な操作を必要としない活動量計を用いた認証手法が提案されている. しかし, 市場には多種多様な活動量計が販売されており, 機種ごとに認証手法を作成することは困難である. 活動量計に依存しない認証手法の研究が必要となるが, 先行研究では1種類の活動量計に対して単一の活動量尺度を用いたモデルの有効性が示されているだけである. そこで本稿では, 先行研究の認証モデルが活動量計の種類, 活動量の尺度によらず有効な手法かを評価した. 評価実験では, 先行研究のFuelBandを用いた被験者70人の活動履歴データ42日分よりFRR 11.51%, FAR 12.03% (活動量の尺度: Fuel)の結果を得た. 2種類の活動量計JawboneとFitbitを用いて新たにデータ収集を行い, 被験者14人の各30日分の活動履歴データから, JawboneでFRR 10.00%, FAR 12.31%, FitbitでFRR 9.22%, FAR 12.92% (いずれも活動量の尺度: 歩数)という結果を得た. 3種類の活動量計を用い, 先行研究の70人の被験者データでの実験に加え, 新たに2種類の活動量計を用い同一の被験者14人で実験することで, 他種の活動量計や他の活動量の尺度でも利用可能な認証モデルであることを確認した.

キーワード: 行動認証, 活動量計, ウェアラブル端末, 機械学習

Evaluation of Activity Behavioral Authentication Model by Comparison between Three Activity Trackers

HIROYA SUSUKI^{1,a)} RIE SHIGETOMI YAMAGUCHI^{1,b)}

Received: June 12, 2017, Accepted: December 8, 2017

Abstract: The growing popularity of multi-factor authentication necessitates a more easy-to-use approach. The more authentication factors the system uses, the more difficult it is to use the system. We considered behavioral authentication as one of the authentication factors. This authentication factor can make the system easy to use by taking advantage of unconscious user behavior. In this paper, we propose a behavioral authentication model that can adapt to various activity trackers for general use using machine learning. The motivation for using an activity tracker as behavioral authentication is that the user simply needs to wear an activity tracker for authentication. However, existing studies are applicable only to the specific activity tracker. For the evaluation of our model, we compared the difference between activity measures, such as Fuel and steps, and the difference between three activity trackers. The results showed that it is possible to apply our model to various activity measures and various activity trackers. Our model achieved a false rejection rate (FRR) of 10.83% and a false acceptance rate (FAR) of 10.30% with the FuelBand activity tracker, an FRR of 10.00% and an FAR of 12.31% with Jawbone, and an FRR of 9.22% and an FAR of 12.92% with Fitbit.

Keywords: behavioral authentication, activity tracker, wearable device, machine learning

1. はじめに

近年, デジタルデータが飛躍的に増大しており, 大規模データを扱ったビッグデータ活用技術に注目が集まってい

¹ 東京大学
The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan
a) susuki.hiroya@sict.i.u-tokyo.ac.jp
b) yamaguchi.rie@i.u-tokyo.ac.jp

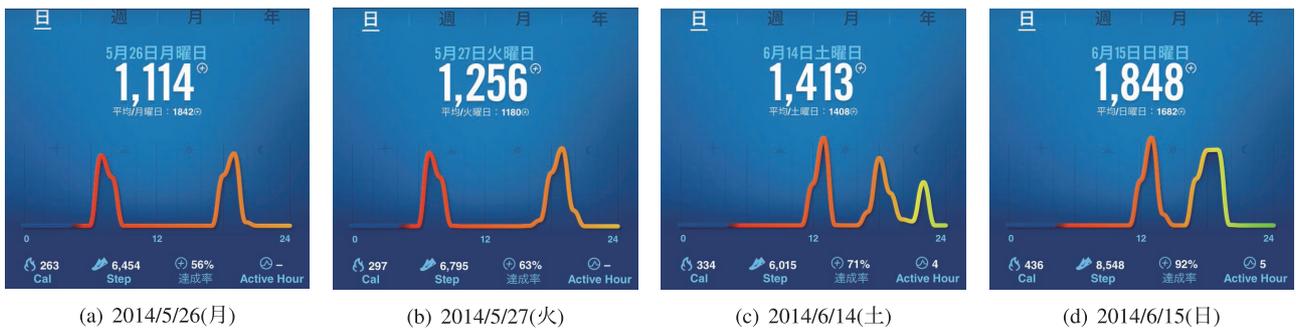


図 1 FuelBand で収集した 4 日分の 24 時間ごとの活動量グラフ。縦軸：活動量 (Fuel)，横軸：時間 (時)

Fig. 1 Four graphs of activity data collected every 24 hours in the FuelBand. Vertical axis: activity (Fuel). Horizontal axis: hours.

る。ビッグデータの活用では、人手の解析では処理しきれない大規模データに対して機械学習技術を用いることで、従来の手法では得られなかった有益な情報を得ることができる。ビッグデータは製品の品質管理や人間の行動分析、マーケティング、レコメンデーションなどに活用されている。応用技術の 1 つであるレコメンデーションは、多様なデータからユーザの行動履歴を用い、各個人の特徴や趣味嗜好に合わせたサービスを提供する技術である。多量のデータを収集できるようになったことで、このように行動履歴データからユーザ個人の特徴を見つけることまでできるようになった。従来の行動履歴は収集可能なデータ量が少なかったが、個人を識別するだけの十分な情報量を得られるようになり、個人認証も応用先の 1 つとして考えられている。

スマートフォンの普及により、日常の様々な場面でオンラインサービスを利用するために認証が行われており、個人認証技術の重要性は年々増している。一方で、導入コストの低さとユーザの利便性の観点から、いまだ多くのサービスでパスワード認証が利用され続けており、安全性の低下が問題となっている [1], [2]。この問題を解決するために、パスワードと指紋認証など複数の認証要素を組み合わせる多要素認証が提案されている [3], [4]。認証要素の増加は安全性の向上につながる一方、読み取り機に指をかざす動作が追加されるなど利便性が低下するという課題もある。この課題を解決するため、行動的特徴を用いた認証 (以後、行動認証と記す) の研究に注目が集まっている [5], [6]。

1.1 活動量を用いた行動認証の背景と目的

行動認証の例としては、歩容認証 [7] や署名認証 [8]、キーストローク [9] やタッチ操作 [5] を用いた認証がある。いずれも、人間の歩行動作や、署名動作における筆跡や筆圧、キーボード操作時のストロークの特徴、タッチスクリーンの操作における指運の特徴など、時系列の履歴情報を用いている。これらの行動認証と同様に提案されている手法として、活動量計を用いた認証手法 (以後、活動量認証と記す) がある [10]。

活動量は時系列で収集可能な行動履歴データの一種である。活動量とは日常生活で消費するエネルギーを数値化したもので、従来は歩数計が計測手段の 1 つとして使われていた。これまでの歩数計は合計歩数のみが記録されていたが、近年ではスマートフォンや市販の活動量計などを用いて、歩数や消費カロリーを時系列に記録することができるようになってきている。時系列の活動量データを得られるようになったことで、実際に個人識別に活用できるという事例を図 1 により示す。図 1 の各図はそれぞれ同一人物の 4 日分の活動量を、後述する FuelBand を用いて収集し、示したものである。いずれも移動中のみ活動量計を着けた場合の 24 時間の活動量の推移を、活動量計のアプリで表示している。図 1(a) と図 1(b) は、同じ場所に同じ通勤経路で勤務地に移動した 5/26 (月)、5/27 (火) の両日の結果を示しており、1 日の活動量を示す波形が類似していることが確認できる。また、図 1(c) と図 1(d) は、同じ場所 (5/26, 5/27 とは別の場所) に、行きのみ同じ交通経路で移動した 6/14 (土)、6/15 (日) の両日の結果を示している。こちらは両日の 1 つ目の山の波形が類似していることは前と同様で、異なる経路で帰った 2 つ目以降の山の形状が異なることが確認できる。これらのデータは、人間にとっては直感的に同じ人間の行動であると推測することができる。

このような特徴と近年の活動量計の普及に着目し、著者らが新たに提案したのが活動量計を用いた認証手法である [10]。従来、活動量に類似した人間の動きの情報を活用した認証手法として、加速度センサを用いた歩容認証の研究が行われてきた [11], [12]。歩容認証は、人間の歩く動作に含まれる個人ごとの特徴を用いた認証手法である。加速度センサを用いた歩容認証も行動認証の一種であるが、歩行動作以外では認証に適用できないという問題がある。さらに、認証のために加速度センサを身につける必要があり、利便性の観点でも問題があった。活動量認証は、近年の健康志向の高まりによりヘルスケアやフィットネス用途で利用者が増加している市販の活動量計を認証に用いるこ

とで、認証専用のセンサを身につける必要がない点に特徴がある。また、常時収集可能な活動量を用いることで、歩行時以外の行動特徴も認証に活用できる。

著者らが提案した活動量計を用いた認証と同様の研究報告は他になく、新規性の高い手法である。しかし、著者らの先行研究では単一の活動量計に対し、単一の活動量の尺度だけを用いて提案手法を評価しており、他の種類の活動量計に対しても有効かどうかは今後の課題としていた。総務省の資料によると、活動量計などスマートバンドに分類されるウェアラブルデバイスの市場規模は2015年の時点で約140万台あり、2020年までには350万台に達すると予測されている[13]。2017年9月時点でヨドバシカメラのオンラインショップで「活動量計」として取り扱われている商品だけでも626件あり、多種多様な活動量計が販売されていることが分かる。これらの活動量計1つずつについて最適な認証手法を研究することは、量の観点で現実的ではない。そもそも個々の活動量計に限定して研究を行うことは、特定の端末でしか使えない認証手法となり、利用場面の限定につながるという問題がある。たとえば、ある企業が自社社員の認証のために同一機種の活動量計を配り利用するという用途では有効である。一方で、認証用途以外の目的で所持している端末からデータを取得することでユーザ負荷を下げられる、という活動量計認証の利点を消すことになってしまう。さらに、各活動量計が収集する活動量の尺度には、歩数やカロリー、METsといった一般的な単位からNike社が独自に定義しているFuelといった単位まで様々な単位がある。このような背景から活動量計の種類、活動量の尺度によらず適用可能な手法の研究が求められている。

また今後、異なる活動量計や異なる活動量の尺度で認証手法が提案された場合、個別の手法だけで他との差を比較することは困難であり、共通の基準を示す必要がある。先行研究で提案した著者らの手法は、活動量を時系列の数値情報としてのみ扱っている。活動量計の種類、活動量の尺度に依存する情報を用いておらず、基準となりうる手法である。本稿では、3種類の活動量計と2種類の活動量の尺度に適用することにより、提案手法が特定の活動量計にしばられない手法かどうかの評価を行った。

1.2 本研究の目的と論文構成

本研究の目的は、先行研究で提案されている活動量計を用いた認証手法が特定の活動量計にのみ有効かどうかを評価し、認証モデルの有効性を確認することである。著者らの貢献は、まず先行研究の認証モデルを整理し、活動量計によらないモデルとして定義し活動量から抽出する特徴量を示したことである。さらに、このモデルに対して先行研究と同じデータを用い、Nike社の活動量計Nike+ FuelBand SE[14]（以後、FuelBandと記す）を用いて被験者70人か



図2 活動量計：Nike+ FuelBand SE (左上)、UP24 by Jawbone (中央)、Fitbit Flex (右下)

Fig. 2 Photo of activity trackers: Nike+ FuelBand SE (top left), UP24 by Jawbone (center), Fitbit Flex (bottom right).

ら得た42日分の活動量データで、本人拒否率(FRR)と他人受入率(FAR)により認証精度の評価を行った。先行研究では、評価に正解率(Accuracy)のみを示しており、FRRとFARについては評価されておらず、本研究で改めて算出したものである。さらに、提案モデルが特定の活動量計や特定の活動量の尺度に依存するかを評価するために、Jawbone社のUP24 by Jawbone（以後、Jawboneと記す）[15]とFitbit社のFitbit Flex（以後、Fitbitと記す）[16]の2種の活動量計を用いた実験を行った。2種の実験を同じ被験者で実施することで、活動量計認証モデルが活動量計によらない特徴を持つことを評価した。後者の実験では、JawboneとFitbitの双方を同一の14人の被験者に身につけてもらい30日分ずつのデータを収集し、モデルの有効性について解析を行った。両活動量計は2015年のフィットネストラッカ市場のシェア1, 2のメーカーが販売しているものを用いており、広く知られた製品である[17]。さらに、先行研究との被験者数の違いの影響を考慮し、FuelBandの実験についても人数を揃えて再実験を行い、Jawbone、Fitbitと同じ14人ごとの認証精度を求めた。なお、FuelBandは販売を終了しており、比較のため3種を同一の被験者で評価する実験はできなかった。図2は3種の活動量計の写真である。

本稿の構成は次のようになっている。2章では、行動認証手法および行動認証に用いられるデバイスについて紹介する。3章では、活動量計を用いた行動認証手法と活動量の特徴量について説明する。4章では、実際に被験者から収集したデータをもとに、機械学習を用いた提案システムによる認証の実験結果を示し、5章で考察を行い、6章で結論を述べる。

2. 関連研究

2章では、関連研究として行動認証手法について紹介し、歩容認証、活動量認証などの行動認証に用いられるデバイスについても述べる。

2.1 行動認証

個人認証の研究では安全性に着目した手法の提案が多くなされているが、実社会での利用という観点において、利便性の考慮も重要な要素である。

行動認証は、行動の中に無意識的に含まれる本人性を利用することで、ユーザの負荷を下げるができる。行動認証が実際に利用されている事例の1つとしてリスクベース認証がある。Googleはアクセス元IPアドレスの履歴などを用いたリスクベース認証を行っている[18]。一般に、個人がグローバルIPアドレスを占有していることはないため、IPアドレスには単独で個人を特定できるだけの一意識別性はない。しかし、たとえばふだんは日本のIPアドレスからアクセスされているユーザのアカウントに対して、海外から認証要求があった場合には高リスクと判断して追加の認証を要求したりすることができる。これは行動認証がユーザの行動におけるふだんの規則性をもとに認証を行っているという特徴に依拠している。しかも、ユーザの認証操作そのものに何かを追加する必要がないという利点がある。このように、行動認証は一意識別性が低くても、ユーザ自身が特別な操作を行う必要がない点に有用性がある。この行動認証の有用性は、多要素認証の一要素として利用する場合も同様であり、認証要素の増加とともに低下する傾向にある利便性を損なわずに利用できる。

2.2 歩容認証と活動量認証の違い

活動量認証と同様に、人間の身体の動きを用いて明示的な入力操作を必要としない手法として、歩行動作を行動的特徴とした歩容認証がある。歩容認証には、歩行動作の画像を利用する手法[7], [19]と、加速度センサを利用した手法[11], [12]の2種類がある。歩行動作の画像データを取得するためにはカメラによる撮影が必要であり、カメラが設置されていない場所での認証が行えないという問題がある。

加速度センサを用いた歩容認証は、3軸の加速度データから個人の特徴を抽出することで認証を行っている[12], [20]。これらは廊下などの特定区間を歩行する際の特徴に着目しており、限られた歩行区間で認証に必要なデータ量を得るため、ms単位などの短いサンプリング間隔で収集したデータを用いている。一方で、サンプリング間隔が短いセンサはバッテリーの消費も大きく、常時センサを所持しなければならない日常的な認証での利用には適さない。活動量計はヘルスケアなどを目的においており、サンプリング間隔が長い代わりに数日から1カ月以上データを収集し続けることができる。また、加速度センサを用いた歩容認証においては、歩行時の特徴のみが使われているため、歩行以外の状況での認証に適用できないという問題もある。個人認証の適用場面を考えた場合、歩容認証は入退室管理や防犯目的では有効であるが、日常的な場面で認証に用いるには場

面が限定される。

このように、活動量計を認証に用いる利点として、歩容認証と同様に認証時に明示的な操作が不要であり、かつ、高精度の加速度センサのような専用デバイスではなく、誰もが利用できる市販の製品をデバイスとして使えるという点がある。

2.3 活動量認証に利用可能なデバイス

従来、個人認証には様々なデバイスが用いられており、行動認証では特にウェアラブルデバイスが活用されている。現在、ウェアラブルデバイスには様々な用途のものがあり、その形状も多様である。主に、腕輪（腕時計）型、指輪型、眼鏡型、アクセサリ型などがあり、各デバイスの用途に合わせた機能やセンサ類を内蔵している。活動量の計測を主な用途としているウェアラブルデバイスには腕輪型が多く、FuelBand[14]やJAWBONE[15]、Fitbit[16]などがある。機器ごとに差異はあるが、多くの活動量計はカロリーや歩数、METsなど広く一般に用いられている活動量の尺度で活動履歴データを記録する。

ウェアラブルデバイス以外に、加速度センサを備えた機器としてスマートフォンがある。入力手段の限られるスマートフォンの利便性を考慮した個人認証の研究として、ユーザの位置情報[3]や端末の操作履歴、アプリケーションの利用状況[21]などの履歴情報を組み合わせて多要素認証を行う手法が提案されている。一方で、Rivaらが論じているように、スマートフォンは手で持ったり、机に置いたりするため、つねに所持しているわけではない[22]。そのため、身体に身に付けるウェアラブルデバイスと違い、スマートフォンは認証のために常時活動量を収集する用途には適していない。

2.4 活動量計を用いた認証

市販の活動量計を用いた研究には、医療[23]やヘルスケア[24]、フィットネス[25]を目的としたものが多く、個人認証に適用した研究は著者らの研究以外にまだない[10]。著者らの手法は機械学習を用いて分類することで個人認証を行うものである。

機械学習を用いた分類手法には、1対1分類と1対多分類がある。個人認証を行う場合、本人と本人以外に分類するため、AさんとAさん以外を分類するように学習を行う必要がある。分類には単純に1対多分類を用いる手法と、AさんとBさん、AさんとCさんを分類する1対1分類を組み合わせる手法がある。著者らの先行研究では2つの手法を比較しており、1対多分類よりも1対1分類の組合せの方が精度が良いことを報告している。1対多分類では、多クラス分類が可能なSupport Vector Machine (SVM)とRandom Forestの2種類で評価し、1対1の組合せでは1対1分類を行うSVMを用いて評価を行っている。1対多

分類の精度が低い要因として、活動量計から得られる情報量の少なさによるものと考察している。1対1分類ではAさんとBさんの間の違い、AさんとCさんの間の違いを別々に学習することができる一方、1対多分類の学習ではAさんとAさん以外全体の間の違いを学習する必要がある、必要となる情報量が増える。しかし、市販の活動量計は、歩容認証における3軸加速度センサと比べてデータの収集間隔が長く、また3軸加速度が1次元の活動量に換算されることで情報量が減っており、十分な分解能が得られないという制約がある。そのため、1対1分類が適していると結論している。そのうえで、著者らは1対1分類の欠点である組合せ数の増加による学習コストの問題に対して、精度への影響を少なくコストを低減させるための手法を提案した。

このように、著者らの先行研究では機械学習のコストに焦点を当てており、1種類の活動量計でしか実験しておらず、多種の活動量計に有効かどうかは示していない。異なる活動量計や異なる活動量の尺度に適用可能でなければ、特定の活動量計に依存しない認証手法とはいえない。また、同じ尺度を用いる異なる活動量計に適用可能かどうかについても評価されていない。しかし、まだこのような評価を行った研究の事例はない。

従来、活動量認証は活動量計を持った一部の利用者のみ使える認証手法であったが、今後、ますます普及することが予測されており、活動量計を用いた認証を適用可能な条件が揃ってくる。一方で、多くのメーカーから多種の活動量計が販売されており、広く利用するためには特定の活動量計に依存しない認証手法が必要である。

3. 活動量計を用いた行動認証のモデルと活動量データの特徴量

3章では、活動量計を用いた活動量による個人認証のモデルと認証に用いる活動履歴データから抽出する特徴量について述べる。

3.1 活動履歴を用いた認証モデル

先行研究では、2値分類器の組合せを用いた機械学習のコストに着目しており、認証モデルの定義は行っていなかった。そこで本稿では、2値分類器の組合せによる活動量を用いた認証モデルの定義を改めて行った。以降のモデル定義の表記は文献[26]を参考とした。モデルそのものは先行研究と同じである。図3は2値分類器の学習と分類を示した図である。

(i) 学習

事前に2値分類器の関数 f があるとする。 n 個の事例からなる学習データを $\{(x_i, y_i)\}_{i \in [n]}$ としたとき、 $x_i \in \mathbb{R}^d$ を d 次元実数の入力ベクトル、 $y_i \in \{1, -1\}$ を1か-1の値をとるラベルとする。 x と y に添字がない場合は一般的な

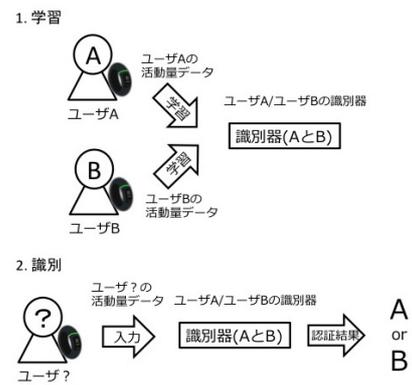


図3 2値分類器による学習と分類
Fig. 3 Learning and classification by binary classifier.

ベクトルとラベルを表す。

α と β はユーザを表す ($\alpha, \beta \in U \mid U$: 認証システムの登録ユーザ集合)。ある2人の登録済みユーザ α, β を分類する識別器を作成する場合、それぞれ α から事前に収集した n_α 個の学習用の活動量データ $\{(x_l, y_l)\}_{l \in [n_\alpha]}$ と、 β から事前に収集した n_β 個の活動量データ $\{(x_m, y_m)\}_{m \in [n_\beta]}$ を足し合わせて学習データ $\{(x_i, y_i)\}_{i \in [n_{\alpha\beta}]}$ ($x_i = x_l \cup x_m, y_i = y_l \cup y_m$) を作成する。 α の学習データ x_l のラベル y_l は1、 β の学習データ x_m のラベル y_m は-1とする。

このとき、学習データ $\{(x_i, y_i)\}_{i \in [n_{\alpha\beta}]}$ を関数 f に学習させた結果として、入力データに対して α か β に分類する分類関数 $f^{\alpha\beta}$ を生成する。

$$f^{\alpha\beta}(x) = \begin{cases} 1 & \alpha \text{ と分類} \\ -1 & \beta \text{ と分類} \end{cases}$$

学習された $f^{\alpha\beta}(x)$ は、入力データ x が (α, β) のデータのどちらのユーザの活動量に分類されるかを返す。

分類結果 $r^{\alpha\beta}$ は、関数 $f^{\alpha\beta}$ の結果として1または-1の値をとる。

$$r^{\alpha\beta} \leftarrow f^{\alpha\beta}(x) \quad (\alpha \in U, \beta \in U \text{ (ただし, } \alpha \text{ を除く)})$$

(ii) 識別

x_{test} (x_{test} : 活動量データ) がユーザ u ($u \in U$) の活動量かどうかを識別する認証関数 F^u を定義する。認証関数 F^u のアルゴリズムを示す。

ここで、閾値 k は *Accept* と *Reject* の識別結果を調整可能なセキュリティパラメータである。また、 N は登録ユーザ数 $\#U$ からユーザ u を除いた数とする。

$$Result \leftarrow F^u(x_{test})$$

Result は、認証関数 F の結果を格納する。

$$Result = \begin{cases} \text{Accept} & u \text{ として認証成功} \\ \text{Reject} & u \text{ として認証失敗} \end{cases}$$

ここで、 *Accept* とは入力された活動量データ x_{test} が、ユー

表 1 活動量を用いた認証の機械学習に用いる特徴量

Table 1 Machine learning features for activity authentication.

	特徴の種類	詳細
(1)	活動量	生データ (1日 1,440 プロット)
(2)	曜日	データ収集日の曜日情報 (月, 火, 水, 木, 金, 土, 日)
(3)	合計活動量	1分ごとのデータの揺らぎを吸収するため単位時間ごとに合算 (60分, 30分, 10分)
(4)	24時間ヒストグラム	1日に観測された活動量のヒストグラム
(5)	1時間ごとのヒストグラム	0時から23時まで各1時間ごとに観測された活動量のヒストグラム (1日あたり 24組)
(6)	ヒストグラムの変化量	24時間ヒストグラムの隣り合う級での変化量
(7)	活動量の変化量	前後の時間での活動量の変化 (1分ごと, 2分ごと)

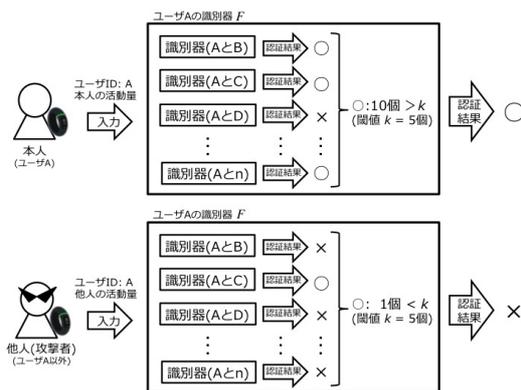


図 4 2値分類器を組み合わせた認証モデル. 本人と他人が認証を実行した場合の例 (閾値 k を 5 にした場合)

Fig. 4 Ensemble learning model using binary classifiers. The example shows two types of authentication results (the threshold value k is 5).

ザ u の活動量データとしていると識別された場合を表し, $Reject$ とは u のデータの活動量データではないと識別された場合を表す.

図 4 は関数の学習と識別を示したものである. 認証システムに N 人のユーザが登録されている場合, 本システムでは 1 人に対して $N - 1$ 個の識別器を用い, 各識別器の識別結果 r をもとに最終的な認証結果 $Result$ を返すモデルである.

Algorithm 1 認証関数 F^u

```

count ← 0
for j = 1 to N do
  if  $r^{uj} = 1$  then
    count ++
  end if
end for
if count ≥ k then
  return Accept
else
  return Reject
end if
    
```

3.2 活動量データから取り出す特徴量

学習と識別の際, 関数 f への入力データとして, 活動量

データを学習データやテストデータに変換して入力する. 学習データとテストデータは, 入力される活動量データそのものではなく, 入力データから活動量の特徴を取り出して作成したデータを用いる.

本研究で, 認証時に入力データとして利用するのは, 1 日分の活動量データとそのデータを収集した日付情報のみである. 入力された活動量データと日付情報から表 1 で示した (1) から (7) の特徴量データを得る.

(1) は 1 日分の 1 分ごとの活動量データ ($X = x_1, x_2, x_3, \dots, x_{1440}$) そのものである. (2) は活動量データの日付から算出し, 月曜から日曜までの各曜日を数値 (1, 2, 3, ..., 7) に置き換えたものである. (3) は (1) の 1 日分の活動量データ (X) を, 0 時 00 分から 23 時 59 分まで先頭から 60 分区切り ($sum60_1, sum60_2, \dots, sum60_{24}$), 30 分区切り ($sum30_1, sum30_2, \dots, sum30_{48}$), 10 分区切り ($sum10_1, sum10_2, \dots, sum10_{144}$) で合算した連続値である. たとえば, $sum60_1$ は $sum60_1 = x_1 + x_2 + x_3 + \dots + x_{59}$ で表される. (4), (5) は, (1) の活動量をヒストグラムにしたものである. 級の区間は活動量の 1 単位分とし, FuelBand であれば 1Fuel 間隔, Jawbone と Fitbit は 1 歩間隔で, 0 から観測された活動量の最大値までの区間の度数を求めた. (4) は x_1 から x_{1440} のすべてからヒストグラムを作成し, (5) は 1 時間ごとに x_1 から x_{59} までのヒストグラムと x_{60} から x_{119} までのヒストグラムというように, 24 個分のヒストグラムを作成した. 特徴量には, 各級の区間の度数 ($h_1, h_2, h_3, \dots, h_n$) を用いた. n は活動量の最大値である. (6) は (4) のヒストグラムの変化量 ($\Delta h_1, \Delta h_2, \Delta h_3, \dots, \Delta h_{n-1}$) を用いた. Δh_1 は $\Delta h_1 = |h_2 - h_1|$ のように表される. (7) は (1) の 1 分ごとの変化量 ($\Delta x_1, \Delta x_2, \Delta x_3, \dots, \Delta x_{1439}$) と, 2 分ごとの変化量 ($\Delta 2x_1, \Delta 2x_2, \Delta 2x_3, \dots, \Delta 2x_{1438}$) である. Δx_1 は $\Delta x_1 = |x_2 - x_1|$, $\Delta 2x_1$ は $\Delta 2x_1 = |x_3 - x_1|$ のように表される連続値である. これらの特徴量は特定の活動量計や特定の活動量の尺度に依存しない値である.

4. 活動量計を用いた認証の評価実験

4 章では, 複数の活動量計を用いて認証モデルの有効性について評価を行う. まず, 認証モデルの実装に用いた機

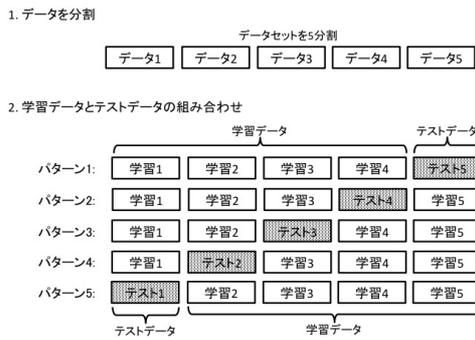


図 5 交差検定の学習データとテストデータ (例: 5 分割)

Fig. 5 Training data and test data of cross validation (5-fold cross validation).

機械学習手法について説明し, FuelBand, Jawbone, Fitbit のそれぞれを用いた実験について実験方法と結果を示す.

4.1 本実験における機械学習

本研究では, 識別器として機械学習手法の 1 つであるサポートベクタマシン (SVM) を用いた. SVM は教師あり学習を行う 2 値分類器であり, 本実験ではその実装の 1 つである LIBSVM [27] を利用した.

SVM の学習に用いるカーネルには RBF カーネルを使用し, 10 分割交差検定で学習を行い, その分類精度をもとに評価を行った. 交差検定とは, 収集したデータを cv 個に分割し, 1 個をテストデータ, 残り $cv - 1$ 個を学習データとして, テストデータと学習データを入れ替えながら cv 回評価を行う手法である. 図 5 は 5 分割交差検定を示した図である. 交差検定には, テストデータと学習データを入れ替えながら学習と評価を行うことで少ないデータでもデータセットの偏りによる誤差が小さくなるという特徴があり, 機械学習を行う場合に広く用いられている手法である. また, 学習パラメータの調整には LIBSVM に実装されているグリッドサーチを用いた [28].

各被験者から収集した活動量の履歴データは 1 日分のデータを 1 件とした. 活動量計は 1 分間隔で活動量データを記録しており, 1 時間ごとに 60 プロット分, 1 日 (24 時間) あたり 1,440 プロット分のデータが得られる. また, 今回の実験では日の区切りを 0 時 00 分とし, 0 時 00 分から 23 時 59 分までを 1 日とした.

4.2 3 種の活動量計を用いた評価実験

ここでは, 3 章でモデル化した認証モデルを用い, 2 つの実験を行った. 各実験の目的とそれぞれの実験に用いた活動量計について示す.

FuelBand を用いた実験 (#1) 先行研究との比較のため, 先行研究と同じデータを用いて評価を行う.

Jawbone と Fitbit を用いた実験 (#2) 先行研究と異なる活動量計を用いた場合の評価を行う. また同じ被

表 2 実験に用いたデータの条件

Table 2 Experimental condition for data collection.

実験番号 #	活動量計	活動量の尺度	被験者数 (人)	実験日数 (日)
#1	FuelBand	Fuel	70	42
#2	Jawbone Fitbit	歩数 歩数	14 14	30 30



図 6 FuelBand と FuelBand アプリケーション (iPhone 版)

Fig. 6 Photo of FuelBand and FuelBand application on iPhone.

験者が各活動量計を身につけた場合の違いも評価する.

表 2 は 2 つの実験の実験条件を示しており, 各実験で用いた活動量計とその活動量の尺度, 被験者数と収集データ数について示している. 本稿では 3 種の活動量計を用いて延べ 84 人の被験者から収集したデータを用いた. いずれの活動量計も市販の腕輪型ウェアラブル端末で左右いずれかの腕に装着する (写真: 図 2). 各活動量計とも 3 軸の加速度センサを備えており, 1 分に 1 回の間隔で人間の活動量データを記録する. 記録可能な活動量の尺度には, 歩数, カロリー, METs, Fuel (Fuel は 3 種のうち FuelBand のみ) がある. 本実験の目的は, 今回の実験で収集したデータから活動量計の違いと活動量の尺度の違いが活動量計を用いた認証に与える影響について評価を行うことである. いずれも学習には表 1 で示した特徴量を用いて学習を行う.

4.2.1 FuelBand を用いた実験 (#1)

本実験では, 先行研究で Nike+ FuelBand SE (図 6) を用いて収集されたデータにより実験を行った. FuelBand は, Fuel と呼ばれる活動量の尺度を用い, 1 分ごとの活動量データを記録する. 被験者は, 20 代から 60 代の男女 70 人 (男性 46 人, 女性 24 人) で, 大学生が 25 人, 学生以外が 45 人からなる. 2014 年 5 月 26 日 (月) から 2014 年 7 月 6 日 (日) の 6 週間で実施し, 各被験者から 42 日分のデータを収集したものである. 1 日の着用時間については, 24 時間常時ではなく, 就寝, 入浴時などは任意で取り外すことが許されている.

本実験では, 70 人の被験者に対してそれぞれ識別器を作成した. 図 4 のとおり, 各被験者に対して残り 69 人との

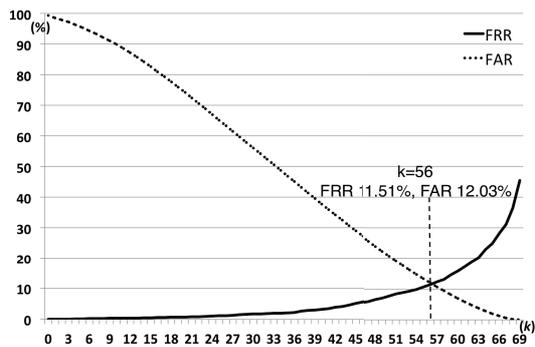


図 7 FuelBand を用い 70 人の被験者から収集した各 42 日分のデータで認証精度を評価。閾値 k を変えた場合の FRR と FAR の変化を示す。縦軸：FRR (実線) と FAR (破線) (%)。横軸：閾値 k

Fig. 7 Experimental result (FRR and FAR) by using FuelBand. The activity data from 70 participants who worn FuelBand during 42 days. Vertical axis: FRR and FAR (%). Horizontal axis: threshold, k .

識別を行うように学習させた識別器を用意し、69 個の識別器の識別結果の合計が閾値 k 以上か、 k 未満かで、本人として受け入れるか拒否するかの認証を行う。1 個の識別器で 2 人分のデータを学習するが組合せの順序は問わないため、作成した識別器の総数は $70 \times 69 \div 2 = 2,415$ となる。10 分割交差検定により、5 月 26 日 (月) から 7 月 6 日 (日) まで日付順に 4 日分ずつ分割して組みを作り、ある組みを識別用のテストデータ、残り 9 組すべてを学習データという形で入れ替えながら 10 回実施し、すべての組みのデータが 1 回のテストと 9 回の学習に使われるように実験している。

図 7 は、FuelBand を用い、先行研究と同じデータで行った実験の結果を示している。先行研究では Accuracy で評価を行っていたが、本稿では、認証モデルの精度を評価するために FRR と FAR を用いる。また、結果として示す値は、閾値 k を変えていき、FRR と FAR の差が最小となる値とした。FuelBand を用いた実験の結果は FRR が 11.51%、FAR が 12.03% となった。

4.2.2 Jawbone と Fitbit を用いた実験 (#2)

FuelBand 実験と同様に、Jawbone (図 8) と Fitbit (図 10) を用いて実験を行った。双方とも FuelBand と同様に 1 分ごとの活動量を記録する。Jawbone と Fitbit では活動量の尺度として歩数を用いて評価を行った。

本実験では、FuelBand と異なる被験者 14 人のデータを収集した。被験者は、20 代から 30 代の大学生 (院生を含む) で男女 14 人 (男性 12 人、女性 2 人) からなる。2015 年 7 月 17 日 (金) から 2015 年 11 月 16 日 (月) の期間で収集した。各被験者には Jawbone と Fitbit それぞれの活動量計を 30 日ずつ身につけてもらった。10 分割交差検定により、3 日分ずつ分割して順に組みを作り、ある組みを識別用のテストデータ、残り 9 組すべてを学習データとい



図 8 Jawbone と Jawbone アプリケーション (iPhone 版)
Fig. 8 Photo of Jawbone and Jawbone application on iPhone.

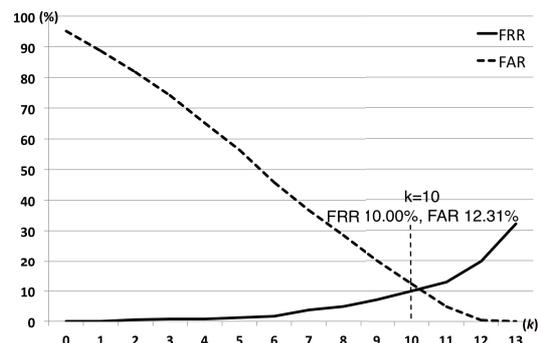


図 9 Jawbone を用い 14 人の被験者から収集した各 30 日分のデータで認証精度を評価。閾値 k を変えた場合の FRR と FAR の変化を示す。縦軸：FRR (実線) と FAR (破線) (%)。横軸：閾値 k

Fig. 9 Experimental result (FRR and FAR) that 14 participants worn Jawbone during 30 days. Vertical axis: FRR and FAR (%). Horizontal axis: threshold, k .



図 10 Fitbit と Fitbit アプリケーション (iPhone 版)
Fig. 10 Photo of Fitbit and Fitbit application on iPhone.

う形で入れ替えながら 10 回実施し、すべての組みのデータが 1 回のテストと 9 回の学習に使われるように実験している。ただし、FuelBand の実験データと違い、全被験者で日付を揃えた実験は行っていない。そのため、日付が重複する被験者もいるが、被験者ごとに各 30 日分のデータの日付は異なっており、10 分割する際は古い日付から日付順に分割している。その他の条件は FuelBand を用いた実験 (#1) と同じである。

表 3 実験条件を揃えた追加実験の条件とその結果

Table 3 Additional experimental condition using extracted data and the experimental result.

実験番号 #	活動量計	活動量の尺度	被験者数	実験日数	FRR	FAR
#1	FuelBand	Fuel	70	42	11.51%	12.03%
#1.2	FuelBand	Fuel	14 × 5	30	10.83%	10.30%
#2	Jawbone	歩数	14	30	10.00%	12.31%
	Fitbit	歩数	14	30	9.22%	12.92%

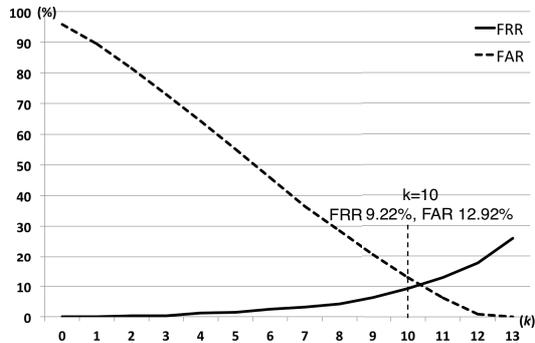


図 11 Fitbit を用い 14 人の被験者から収集した各 30 日分のデータで認証精度を評価。閾値 k を変えた場合の FRR と FAR の変化を示す。縦軸：FRR (実線) と FAR (破線) (%)。横軸：閾値 k

Fig. 11 Experimental result (FRR and FAR) that 14 participants worn Fitbit during 30 days. Vertical axis: FRR and FAR (%). Horizontal axis: threshold, k .

実験結果は、Jawbone で FRR 10.00%, FAR 12.31% (図 9), Fitbit で FRR 9.22%, FAR 12.92% (図 11) となった。Jawbone と Fitbit それぞれから得た 14 人の被験者の FRR と FAR に有意な差があるか t 検定を用いて評価を行った。有意水準 5% で対応がある標本に対する両側検定を行い、FRR の p 値は 0.593, FAR の p 値は 0.837 となった。いずれの p 値も 0.05 より大きく、 t 検定の結果からそれぞれの FRR, FAR に有意な差があるとはいえないという結果となった。同じ被験者群から収集した活動量データに対して同じ認証モデルを適用した結果、同程度の FRR, FAR を示しており、提案モデルが活動量計の違いに有意な影響を受けていないことが分かった。

4.3 FuelBand と Jawbone, Fitbit の比較

本節では、FuelBand の実験 (#1) と Jawbone と Fitbit の実験 (#2) の 2 つの実験結果を比較する。異なる活動量計で異なる活動量の尺度を用いたそれぞれの実験を比較するため、被験者に関する実験条件を調整した。表 2 のとおり、FuelBand の実験 (#1) と Jawbone と Fitbit の実験 (#2) では被験者数とデータの収集日数が異なっている。この違いを揃えるため、FuelBand の実験 (#1) の被験者データから 30 日分を抽出した。元データのデータ収集期間である 2014 年 5 月 26 日 (月) から 2014 年 7 月 6 日

(日) のうち、2014 年 5 月 26 日 (月) から 2014 年 6 月 30 日 (月) までの 30 日分を使用した。また、被験者数 70 人をランダムに 14 人ずつの 5 組に分けた。表 3 は、被験者数とデータの収集日数の実験条件を揃えた場合の実験結果を示している。FuelBand の実験 (#1.2) の FRR と FAR は 5 組の平均であり、FRR 10.83%, FAR 10.30% となった。この結果は、70 人 42 日の実験 (#1) より FRR, FAR とともに良い結果となっている。これは被験者数を揃えたことで他人データが少なくなり FAR が向上し、実験日数を減らしたことで日によるばらつきの影響が少なくなったためと考えられる。Jawbone と Fitbit を用いた実験 (#2) の結果と比較すると、FAR は 2% ほど低いが、FRR は同じ程度の値を示している。

5. 考察

5 章では、活動量計ごとの違いや活動量の尺度の違いが提案モデルに与える影響について考察する。

5.1 時間ごとの 1 日の行動習慣

本節では、被験者から収集したデータに 1 日の行動習慣が現れているかについて考察する。図 12 は、FuelBand, Jawbone, Fitbit の時間ごとの平均活動量を示している。FuelBand で収集した活動量を他と同一尺度で比較するため、Fuel は歩数に合わせて正規化している。全体の傾向として、いずれも夜間帯は活動量が小さくなり、日中帯に活動量が大きくなっている。今回の実験データには日常的に夜間作業をするような勤務形態の被験者は含まれておらず、日中に仕事や学業をし、夜間に帰宅し、深夜は睡眠をとるといった一般的な行動習慣が現れている。時系列でデータを見たときに、3 つの山があり、7 時、13 時、18 時の前後の時間で活動量が大きくなっている。いずれも人間の行動習慣にあてはめると、7 時、18 時は通勤通学などで行動により活動量が増える時間帯であり、13 時前後も昼食などで行動による活動量が増える時間帯である。

5.2 収集データの違い

表 4 は、図 12 の活動量が異なる活動量計との間で相関があるか調べた結果である。同一被験者からデータを収集した Jawbone と Fitbit の間の相関係数は、0.993 という値を示しており、高い正の相関が見られる。さらに、

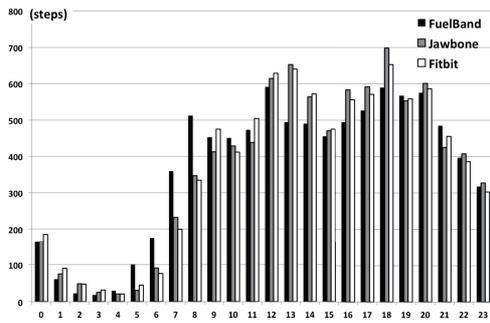


図 12 3種の活動量計の時間ごとの平均活動量を比較 (Fuel は歩数に正規化). 縦軸: 歩数, 横軸: 時間 (時)

Fig. 12 Comparison of hourly average activity, normalized by steps, for three activity trackers. Vertical axis: steps, Horizontal axis: hour.

表 4 時間ごとの平均活動量の相関係数を活動量計の組合せごとに比較. FuelBand は Fuel を歩数で正規化

Table 4 The correlation coefficient of hourly average activity between activity trackers.

活動量計の組合せ	相関係数
FuelBand, Jawbone	0.950
FuelBand, Fitbit	0.949
Jawbone, Fitbit	0.993

FuelBand と他の活動量計を比較した場合も最小で 0.949 となり, 3種の活動量計から得られたデータが高い相関を示していることが分かる. これはいずれの活動量計で収集したデータでも, 1日の活動量を時間ごとに見た場合に同様の傾向を示すことを表している.

3種の活動量計を見比べた場合, FuelBand と他 2つの間の相関に比べ, 相関係数が 0.993 と Jawbone と Fitbit の間の相関が高くなっている. これは Jawbone と Fitbit が同じ活動量の尺度 (歩数) を用いていることと, 同じ被験者群のデータであることによる. FuelBand は異なる尺度 (Fuel) を用いており, データを収集した被験者群も異なっていることが要因と考えられる. 特に, 被験者群の違いに関しては, 被験者の活動量が増加する時間帯に差が現れている. 図 12 の FuelBand のグラフでは, 7時と 8時の活動量が他の 2つの活動量計に比べて高くなっている. 逆に 14時と 15時, 18時は Jawbone と Fitbit が高くなっている. 仮に, 被験者群の違いによる影響がないと仮定し, 活動量計の種類や尺度による違いによる影響と考えた場合, 特定の時間帯だけ活動量が大きくなるということは考え難い. 活動量計は人間の行動の強度に依存してその活動量を記録するものであり, ある活動量計は朝の活動量が高く, ある活動量計は昼に低いということは起こりえないと考えられる. したがって, この違いは異なる被験者群からデータを収集した影響と考えられる.

各実験データの被験者の内訳は, FuelBand が学生 25人と社会人 45人を含むのに対して, Jawbone と Fitbit は 14

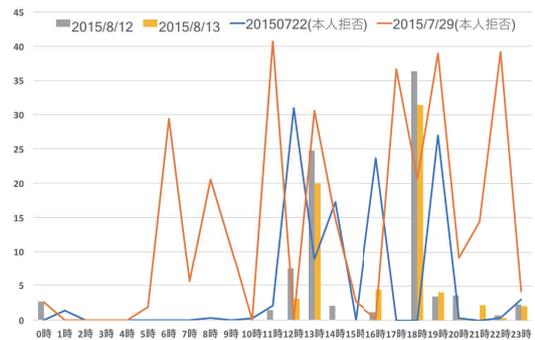


図 13 被験者 ID : 1 の Jawbone で収集した時間ごとの歩数. 折れ線グラフ (2015/07/22 と 2015/07/29) は本人拒否された日, 棒グラフ (2015/08/12 と 2015/08/13) は本人拒否されない日の例. 縦軸: 一時間の歩数 (歩), 横軸: 時間 (時)

Fig. 13 Example of false rejection by collected data from a participant ID: 1 using Jawbone. Line graph shows false rejection data (2015/07/22 and 2015/07/29). Bar graph shows true acceptance data (2015/08/12 and 2015/08/13). Vertical axis: steps. Horizontal axis: hour.

人の学生のみとなっており, 社会人の数の違いが影響を与えていると考えられる. 社会人を多く含む FuelBand では, 毎日決められた勤務開始時間に合わせて規則的に行動する社会人から収集したデータの影響により, 朝 8時をピークとした山ができています. 同様に, お昼休憩の時間が決められていたりする社会人を含む FuelBand は 12時に集中しているのに対して, 受講する授業が曜日によって異なる学生のみ Jawbone と Fitbit は 12時から 14時と活動する時間の幅が広がる傾向が見られる. このように収集した実験データには被験者ごとの行動習慣の違いによる差が見られるが, 一方で, 表 3 のとおり, 実験結果の FRR と FAR は同程度の結果を示しており, 提案手法が学生のみ有効といったように限定的なものではないと考えられる.

5.3 被験者数と本人拒否率と他人受入率の関係

今回の評価の有効性に関して, 被験者数が与える影響を本人拒否率 (FRR) と他人受入率 (FAR) の観点で評価する. Jawbone と Fitbit を用いた実験は被験者数 14人と, FuelBand を用いた実験と被験者群が異なり被験者数も少ない. 70人分と 14人ずつ 5組のデータを用いた FuelBand の実験結果の差から, 被験者数が少ないことで FRR と FAR が低くなっている可能性がある. ただし, 被験者数が与える影響の度合いは FRR と FAR で異なる.

本人拒否率 (FRR) は本人を本人として判定できない確率であり, 本人データがふだんの行動と大きく異なる場合に発生し, 他人データの数に対する影響は他人受入率に比べて小さい. 図 7, 図 9, 図 11 が示すとおり, 閾値 k を変えたときに, FAR が 0%から 100%の幅を持つのに対して, FRR がとる値の幅が小さいことから分かる. 図 13

表 5 FAR が下がるように閾値 k を調整した場合の FRR

Table 5 FRR and FAR when threshold value k minimizes FAR.

活動量計	被験者数	閾値 k	FRR	FAR
FuelBand (#1.2)	14	13	30.99%	0.00%
Jawbone	14	13	31.95%	0.00%
Fitbit	14	13	25.73%	0.00%
FuelBand (#1)	70	69	45.47%	0.00%
	70	68	36.52%	0.13%
	70	67	31.13%	0.61%

は、被験者 ID : 1 の Jawbone の被験者データから 4 日分 (2015/07/22, 2015/07/29, 2015/08/12, 2015/08/13) を示している。被験者 ID : 1 は 30 日のデータのうち 2 日分のデータを除くと、残りのデータはすべての学習器が正しく本人と判定しており、28 日分のデータは閾値 k によらず本人拒否率が 0% となる。図 13 の本人拒否が発生した 2 日と、本人が本人として認証された 2 日をそれぞれ比較すると、この被験者の行動の習慣性として 13 時と 18 時に活動する傾向があることが分かる。一方、認証に失敗した 2 日は行動習慣と大きくずれていることが分かる。他の被験者についても同様に、本人拒否の発生は本人データそのもののふだんの行動との違いに主な要因がある。FuelBand の実験と、Jawbone と Fitbit の実験では被験者群が異なるが、FRR に対する影響は限定的と考えられる。

他人受入率 (FAR) については他人 (B さん) を本人 (A さん) と判定する確率であり、提案モデルでは本人 A さんと B さん、A さんと C さん、A さんと D さんで学習した各分類器が、B さんのデータを A さんと判定する数に依存する。被験者数が増えると似た他人のデータが含まれる可能性が増えるが、同時に似ていない他人も増えることとなるため、B さんデータの誤判定が増えたとしても C さん、D さんが適切に他人と分類されることで、A さん以外のすべての他人データで FAR を求めた結果としては同程度の確率になると考えられる。

閾値を調整し本人拒否率を上げれば FAR を低くすることも可能である。表 5 のとおり、閾値 k を FAR が下がるように変更した場合、FuelBand (#1.2)、Jawbone、Fitbit とともに FAR は 0% になった。代わりに FRR が 30% 前後となり、これは本人が認証した場合におおよそ 10 回中 3 回認証に失敗することを意味する。単独の認証要素として見た場合には高い頻度で本人拒否が発生する。しかし、活動量認証はただ活動量計を身につけているだけで利用可能であり、たとえばスマートフォンのロック解除に活動量認証を用い、活動量認証が失敗した場合のみ PIN 入力を行うようにすれば、他人受入は抑えたままで 7 回は何もせずに解除可能となる。これは実際の利用場面でも利便性が高く有用といえる。なお、FuelBand (#1) で 70 人の場合は FAR が 0% のとき、FRR が 45% と高くなっているが、10 回の認

表 6 Fuel と歩数のデータスパースネスの比較。1 日に観測される 1,440 個のデータに含まれる 0 の出現頻度

Table 6 Data sparseness comparison of Fuel and steps. Frequency of 0 per day.

活動量計	活動量の尺度	0	0 以外	合計
FuelBand	Fuel	1019.6	420.4	1440
Jawbone	歩数	1308.5	131.5	1440
Fitbit	歩数	1257.8	182.2	1440

証に対して 5 回入力が不要でも十分な利便性向上につながる。また、他人受入を多少許容し FRR を同程度の 31% とした場合の FAR は 0.6% となっている。

5.4 活動量認証の精度

活動量認証の精度に関する原理的な限界と精度向上の可能性について論じる。一般的に行動認証は生体認証などと比較し、精度が高くない傾向にある。これは行動認証のばらつきの大さに起因する。生体認証においても、指紋認証で指を読み取り機にかざす角度のばらつきや、顔認証における周辺明るさの違いによるばらつきなどから入力される認証情報に揺らぎが生じる。行動認証の 1 つである活動量認証は、人の歩き方や動き方から生じる活動量を入力に用いるためにより大きな揺らぎが生じ、十分な精度を得ることが難しい。表 3 のとおり、FRR と FAR は 10% 前後の結果となっている。本稿では、先行研究で提案した手法を他の端末、尺度に適用する実験を行ったが、他の特徴量を利用したり、機械学習手法を変更したりすることで認証精度を向上させられる可能性はある。ただし、今回実験に用いた市販の活動量計は 1 分に 1 回の活動量の記録にとどまっており、高い精度を出すためには情報量が少ないという制約がある。しかし、これは今後、より高頻度にデータを収集する端末が販売されることで解消されると考えている。すでに、オムロンヘルスケア社が提供している活動量計には市販品として 1 分ごとに計測する端末 [29] だけでなく、医療機関など向けに 10 秒ごとに計測できる端末 [30] がある。現在はボタン電池 1 個で前者の一般向けの端末が約 6 カ月、後者の医療機関向けの端末は 45 日間と電池寿命に大きな差があるが、バッテリーの小型化、大容量化により解消されうるものである。高頻度に活動量が収集可能となれば、活動量認証の精度の上限を上げることができる。

5.5 活動量の尺度 Fuel と歩数の比較

活動量の尺度の違いについて検討する。本稿では 3 種の活動量計を用い、Fuel と歩数の 2 種類の尺度で実験を行った。表 6 は、1 日に観測される 1,440 プロットのデータに含まれる 0 の出現頻度を示したものである。各活動量計は 1 分ごとの活動量を収集するため、1 日に 1,440 回記録を行う。運動を行わず、活動量が観測されなければ 0

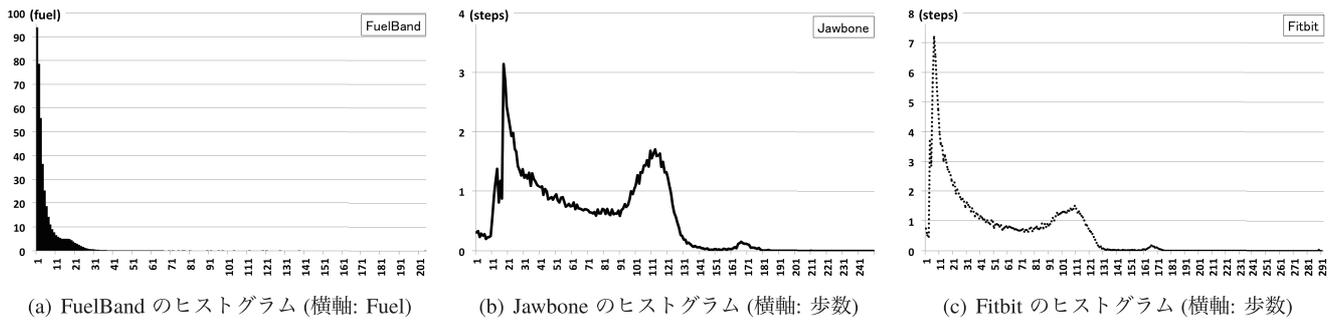


図 14 (a) FuelBand, (b) Jawbone, (c) Fitbit のヒストグラム. 縦軸: 度数, 横軸: 活動量
Fig. 14 Histogram of (a) FuelBand, (b) Jawbone and (c) Fitbit. Vertical axis: Frequency. Horizontal axis: Value of histogram.

が記録される. 表 6 からは Jawbone と Fitbit の歩数と比べ, FuelBand の Fuel は 0 が少ないことが分かる. これは Jawbone と Fitbit が 1 歩に満たない小さな動きの活動量を記録せず切り捨てているのに対して, Fuel はより小さな活動量を記録していることによるものと考えられる. 図 14 は, FuelBand, Jawbone, Fitbit でそれぞれ 1 分ごとの記録された値をヒストグラムにしたものである. 横軸は各活動量の尺度で縦軸がその観測数となる. なお, 図 14 の 3 つのヒストグラムには 0Fuel と 0 歩の値は含めていない. 図 14 を見ると, 1Fuel の観測回数が 90 以上となっており, 小さな活動を多数記録していることが分かる (図 14(a)). 一方の Jawbone と Fitbit は似た傾向を示しており, 1 歩が記録される数は少なくなっている. いずれも 2 つの大きな山があり, Jawbone では 18 と 113 (図 14(b)), Fitbit では 7 と 110 (図 14(c)) をピークとしている. また, 両者とも 165 に小さな山がある. この値が意味することは, 1 分間の行動に 3 つのパターンが多く出現するというを示す. これは実際の行動に対応付けると, ゆっくりとした動きと速く歩く動き, 小走りなどの比較的速い動きの 3 つに相当する.

実際に一般に用いられる徒歩の基準からこれらのピークについて検討してみる. 不動産公正取引協議会連合会が定める「不動産の公正競争規約」第 10 条 (10) [各種施設までの距離又は所要時間] によると, 「徒歩による所要時間は, 道路距離 80 メートルにつき 1 分間を要するものとして算出した数値を表示すること」とある [31]. ここで, 80m に対して歩幅で除算すれば, 1 分間の平均歩数を求めることができる. オムロンヘルスケア社によると歩幅の目安は身長 $\times 0.45$ である*1. 厚生労働省の「国民健康・栄養調査」より, 日本人の成人男女のおおよその平均身長は 170cm と 160cm である [32]. ここでは実験における被験者の男女比の違いを考慮せず, 男女比 1:1 とし, 平均身長 165cm に対して 1 分間に 80m に要する歩数を計算すると 107.7 歩となった. これは Jawbone と Fitbit の収集データにお

*1 http://www.faq.healthcare.omron.co.jp/faq/show/4195?site_domain=jp

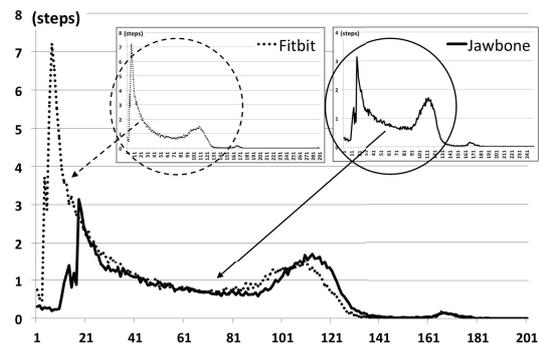


図 15 Jawbone (図 14(b)) と Fitbit (図 14(c)) の歩数ヒストグラムの比較. 縦軸: 度数, 横軸: 歩数
Fig. 15 The Comparison of steps histogram of Jawbone (Fig. 14(b)) and Fitbit (Fig. 14(c)). Vertical axis: Frequency. Horizontal axis: steps of histogram.

ける 2 つ目のピークとほぼ同じ値である. 1 分間歩き続けるという強度の運動は, オフィスや学校, 家庭の中での日常のちょっとした動作で生じるものではなく, 通勤通学や外出など歩行移動の場合に発生するものと考えられる. 3 つ目のピークは小走り程度の少し強い強度の運動と考えられる.

図 15 は図 14 (b) と図 14 (c) を比較するために重ね合わせたものである. 図 14 の両ヒストグラムは同一の被験者が異なる日にそれぞれの活動量計を身につけたデータから得られたものである. Fitbit の低い歩数の度数が大きく記録されていることが分かる. しかし, 14 人の各 30 日のデータの平均として明確な差が現れているということは, 活動量計の差にほかならない. 一方, ヒストグラムの 20 歩以上の度数は Jawbone と Fitbit と類似している. 低歩数すなわち動きが小さい場合に活動量計ごとの差が出る理由としては, 腕輪型のウェアラブル端末の場合, 手の動きから歩数を算出するため, 歩行をとまなう動きか, 歩行をとまなわない動きかの見分けがつきにくく, 各社の算出アルゴリズムの違いが現れているものと考えられる. このような違いがあるが, 提案手法において認証精度の観点で Jawbone と Fitbit に差はなく, 提案モデルが活動量計の差

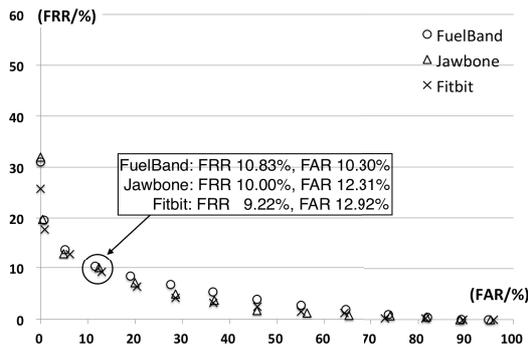


図 16 閾値 k を変えた場合の FRR と FAR : FuelBand, Jawbone, Fitbit. 縦軸 : FRR (%), 横軸 : FAR (%)

Fig. 16 FRR and FAR that are changed by threshold k : FuelBand, Jawbone, Fitbit. Vertical axis: FRR (%). Horizontal axis: FAR (%).

によらず適用可能なことを示している。

このように、活動量計の種類と尺度には違いがあるが、図 16 に示したとおり、FuelBand, Jawbone, Fitbit の FRR と FAR は同程度の値を示しており、提案モデルが活動量計の種類と活動量の尺度に依存しないモデルであることが確認できた。

6. おわりに

本稿では、利便性を考慮した認証手法への要求の高まりに対して、活動量計を用いた行動認証の手法の評価を行った。活動量認証は認証のための明示的な操作を必要としないという利点があるが、一方で多種多様な端末が存在し、活動量の尺度も多数あるため、機種ごとに認証手法を用意することは現実的ではない。著者らは先行研究において提案された認証モデルが他の活動量計や他の活動量の尺度でも利用可能か、実際に 3 種類の活動量計を用いた実験により評価し、その有効性を示した。また、活動量の尺度の違いが与える影響についても評価を行い、先行研究では示されていなかったモデルの有効性について確認した。提案モデルの評価には、先行研究で収集された被験者 70 人の FuelBand の活動履歴データ 42 日分から、FRR 10.83%, FAR 10.30%の結果を得た。さらに、モデルが活動量計の種類と活動量の尺度に依存しないことを評価するために、2 種類の活動量計で被験者 14 人の活動履歴データ 30 日分を収集した。それぞれを提案モデルに適用し、Jawbone で FRR 10.00%, FAR 12.31%, Fitbit で FRR 9.22%, FAR 12.92%の結果を得た。結果から提案モデルが活動量計の持つ違いや活動量の尺度、異なる被験者群に対しても同程度の認証精度を持っていることが確認できた。今後は提案手法で用いていない特徴量や機械学習手法を用いて、活動量認証の精度がどのように変わるかを評価する必要がある。

謝辞 本稿の研究は、次世代個人認証技術講座（三菱 UFJ ニコス寄付講座）による。

参考文献

- [1] 情報処理推進機構：オンライン本人認証方式の実態調査報告書，情報処理推進機構（オンライン），入手先（<https://www.ipa.go.jp/security/fy26/reports/ninsho/>）（参照 2016-02-28）。
- [2] Bonneau, J., Herley, C., Van Oorschot, P.C. and Stajano, F.: The quest to replace passwords: A framework for comparative evaluation of web authentication schemes, *2012 IEEE Symposium on Security and Privacy (SP)*, pp.553–567, IEEE (2012).
- [3] Hayashi, E., Das, S., Amini, S., Hong, J. and Oakley, I.: Casa: context-aware scalable authentication, *Proc. 9th Symposium on Usable Privacy and Security*, p.3, ACM (2013).
- [4] Anzaku, E.T., Sohn, H. and Ro, Y.M.: Multi-Factor Authentication Using Fingerprints and User-Specific Random Projection, *Web Conference (APWEB), 2010 12th International Asia-Pacific*, pp.415–418, IEEE (2010).
- [5] Xu, H., Zhou, Y. and Lyu, M.R.: Towards continuous and passive authentication via touch biometrics: An experimental study on smartphones, *Symposium on Usable Privacy and Security (SOUPS 2014)*, pp.187–198 (2014).
- [6] Zakaria, N.H., Griffiths, D., Brostoff, S. and Yan, J.: Shoulder surfing defence for recall-based graphical passwords, *Proc. 7th Symposium on Usable Privacy and Security*, p.6, ACM (2011).
- [7] 村松大吾, 岩間晴之, 木村卓弘, 横原 靖, 八木康史: 一歩行映像から取得される複数特徴を用いた個人認証, *電子情報通信学会論文誌 A*, Vol.97, No.12, pp.735–748 (2014).
- [8] Yang, Y., Clark, G.D., Lindqvist, J. and Oulasvirta, A.: Free-Form Gesture Authentication in the Wild, *Proc. 2016 CHI Conference on Human Factors in Computing Systems*, pp.3722–3735, ACM (2016).
- [9] Bergadano, F., Gunetti, D. and Picardi, C.: User authentication through keystroke dynamics, *ACM Trans. Information and System Security (TISSEC)*, Vol.5, No.4, pp.367–397 (2002).
- [10] Susuki, H. and Yamaguchi, R.S.: Cost-Effective Modeling for Authentication and its application to Activity Tracker, *The 16th World Conference on Information Security Applications* (2015).
- [11] Nickel, C., Busch, C., Rangarajan, S. and Mobius, M.: Using hidden markov models for accelerometer-based biometric gait recognition, *2011 IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA)*, pp.58–63, IEEE (2011).
- [12] Derawi, M.O., Nickel, C., Bours, P. and Busch, C.: Unobtrusive user-authentication on mobile phones using biometric gait recognition, *2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pp.306–311, IEEE (2010).
- [13] 総務省：情報通信白書平成 28 年版，総務省（オンライン），入手先（<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n3100000.pdf>）（参照 2017-06-11）。
- [14] Nike: Nike+ FuelBand SE, Nike (online), available from (<http://www.nike.com/jp/ja-jp/c/nikeplus-fuelband>) (accessed 2015-03-10).
- [15] JAWBONE, available from (<https://jawbone.com>) (accessed 2016-08-10).
- [16] Fitbit, available from (<http://www.fitbit.com>) (accessed 2016-08-10).

- [17] みずほ情報総研：販売戦略・市場拡大等に関する調査事業，日本医療研究開発機構委託医工連携事業化推進事業（オンライン），入手先 http://www.med-device.jp/pdf/state/summary/AMED2015_marketing_2d_healthcare.pdf（参照 2017-06-11）。
- [18] Google：前回のアカウントアクティビティ，Google（オンライン），入手先 <https://support.google.com/mail/answer/45938?hl=ja>（参照 2016-02-28）。
- [19] Iwama, H., Okumura, M., Makihara, Y. and Yagi, Y.: The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition, *IEEE Trans. Information Forensics and Security*, Vol.7, No.5, pp.1511–1521 (2012).
- [20] Gafurov, D., Helkala, K. and Söndrol, T.: Biometric gait authentication using accelerometer sensor, *Journal of Computers*, Vol.1, No.7, pp.51–59 (2006).
- [21] Khan, H. and Hengartner, U.: Towards application-centric implicit authentication on smartphones, *Proc. 15th Workshop on Mobile Computing Systems and Applications*, p.10, ACM (2014).
- [22] Riva, O., Qin, C., Strauss, K. and Lymberopoulos, D.: Progressive Authentication: Deciding When to Authenticate on Mobile Phones, *USENIX Security Symposium*, pp.301–316 (2012).
- [23] Appelboom, G., Yang, A.H., Christophe, B.R., Bruce, E.M., Slomian, J., Bruyère, O., Bruce, S.S., Zacharia, B.E., Reginster, J.-Y. and Connolly, E.S.: The promise of wearable activity sensors to define patient recovery, *Journal of Clinical Neuroscience*, Vol.21, No.7, pp.1089–1093 (2014).
- [24] Takacs, J., Pollock, C.L., Guenther, J.R., Bahar, M., Napier, C. and Hunt, M.A.: Validation of the Fitbit One activity monitor device during treadmill walking, *Journal of Science and Medicine in Sport*, Vol.17, No.5, pp.496–500 (2014).
- [25] Fritz, T., Huang, E.M., Murphy, G.C. and Zimmermann, T.: Persuasive technology in the real world: A study of long-term use of activity sensing devices for fitness, *Proc. 32nd Annual ACM Conference on Human Factors in Computing Systems*, pp.487–496, ACM (2014).
- [26] 竹内一郎，鳥山昌幸：機械学習プロフェッショナルシリーズ サポートベクトルマシン，講談社 (2015)。
- [27] Chang, C.-C. and Lin, C.-J.: LIBSVM: A library for support vector machines, *ACM Trans. Intelligent Systems and Technology*, Vol.2, pp.27:1–27:27 (online), available from [Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm](http://www.csie.ntu.edu.tw/~cjlin/libsvm) (2011).
- [28] Hsu, C.-W., Chang, C.-C., Lin, C.-J. et al.: A practical guide to support vector classification, available from <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> (2003).
- [29] オムロンヘルスケア株式会社：活動量計 HJ-326J，入手先 http://www.healthcare.omron.co.jp/support/download/manual/pdf/HJ-326F_B_m.pdf。
- [30] オムロンヘルスケア株式会社：活動量計 HJA-750C Active style Pro，入手先 <http://www.healthcare.omron.co.jp/product/hja/hja-750c.html>。
- [31] 不動産公正取引協議会連合会：不動産の公正競争規約，不動産公正取引協議会連合会（オンライン），入手先 <http://www.rftc.jp/kiyak/pdf/kiyak.pdf>（参照 2017-06-09）。
- [32] 厚生労働省：国民健康・栄養調査，厚生労働省（オンライン），入手先 http://www.mhlw.go.jp/bunya/kenkou/kenkou_eiyouchousa.html（参照 2017-06-09）。



鈴木 宏哉（正会員）

報通信学会，言語処理学会各会員。



山口 利恵（正会員）

人工知能学会各会員。

東京大学大学院情報理工学系研究科ソーシャル ICT 研究センター学術支援専門職員。慶應義塾大学工学部情報工学科卒業。同大学大学院修士課程修了。日本サード・パーティ株式会社勤務。2014年5月より現職。電子情報通信学会，言語処理学会各会員。

東京大学大学院情報理工学系研究科ソーシャル ICT 研究センター特任准教授。博士（情報理工学）。産業技術総合研究所研究員，内閣官房情報セキュリティセンター員を経て，2013年6月より現職。電子情報通信学会，人工知能学会各会員。