

## 活動量計データを用いた同行検出手法の研究

## Working-Relationship Detection from Pedometer Data

川尻 亮真<sup>\*1</sup>  
Ryoma Kawajiri坪内 孝太<sup>\*2</sup>  
Kota Tsubouchi荒木 青伊良<sup>\*1</sup>  
Seira Araki下坂 正倫<sup>\*1</sup>  
Masamichi Shimosaka<sup>\*1</sup>東京大学  
The University of TOKYO<sup>\*2</sup>Yahoo! JAPAN 研究所  
Yahoo! JAPAN Research

This paper proposes an innovative way to detect working relationships by using only the step tracking data acquired from pedometers like Fitbit.

The idea makes the cost of working-relationship detection much lower than that of previous approaches. We can find out if people have a working relationship and spend their daily lives together by making them wear a pedometer. Results of an experiment in Japan showed that this approach is very effective and practical. An organizations profile can be written automatically by analyzing the data in correctly.

## 1. 序論

人同士のつながりの強度を示した人間関係グラフは、組織評価やコミュニティ把握において重要な情報となる。人間関係グラフを検出する研究は多くあり、SNS履歴といったウェブ上の行動情報を用いた手法および実世界での情報を用いた手法の2種類に大分される。本稿では組織評価やグループの動的活動の評価を対象としたいため後者を取り扱う。

実世界での情報から人間関係グラフを生成する手法として、たとえば赤外線、Wifiアクセスポイント [Justin 10]、Bluetooth [Jacopo 12] 等の近接センサーやGPS情報を用いる手法が知られている。対象となるユーザのペアが「一緒にいた」回数をカウントし、同行回数を両名の人間関係の強度に置き換える手法である。しかし、これらの手法はデータを取得するために専用の送受信機を身につけなければならない、導入コストが大きい。また、GPS情報やWifiアクセスポイント情報には一緒にいる相手の情報と同時に、その場所までを特定されるためプライバシーの観点からも現実的ではない。

本稿では、入手が用意でかつプライバシーの観点からも障壁の低い万歩計の歩数データからユーザの同行を検出し、人間関係グラフを得る新手法の提案を目的とする。

## 2. 人間関係グラフの生成手法

歩数データから人間関係グラフを生成するために、被験者の同行を検知する。人間関係が強い2名は自然と彼らの生活において同行する回数が増えると想定している。

同行とは、2人で一緒に歩く行為を指す。たとえば、一緒に会議に参加したり、食事をしたり、運動をしたり、といった行為である。通常は互いに一分間の歩数や歩調の異なる2名でも、同行時は互いの歩調を合わせ、また歩き始めや歩き終わりが一致する。この現象に着目し、万歩計データから同行を検出することを試みる。万歩計からは、図1に示す1分あたりの歩数データを得る。

ここで、対象とするユーザ2名の歩数をそれぞれ  $a, b$  とおく。  $a_t, b_t$  はある期間  $t$  における歩数を指す。ここで、  $a, b$  の

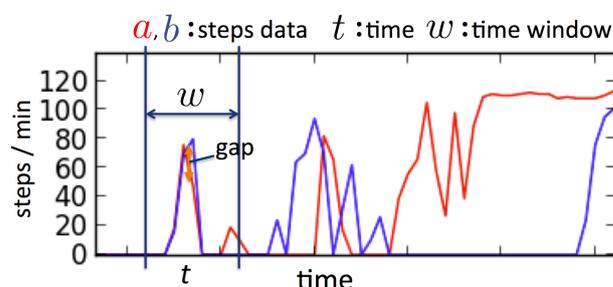


図1: 2名の歩数データの例

歩数差に注目した非類似度  $\Delta_t$  を以下のように定義する。

$$\Delta(a, b)_t = \left( \sum_{\tau=t}^{t+w} (a_\tau - b_\tau)^2 \right) / \alpha_t \quad (1)$$

なお、式中の  $w$  は探索の時間幅を示す。本手法ではある時間幅  $w$  における両者の歩数データの類似度を同行の程度と仮定している。なお、  $\alpha_t$  は、歩数差を正規化するための値で次の通り定義する。

$$\alpha(a, b)_t = \sum_{\tau=t}^{t+w} (a_\tau^2 + b_\tau^2) \quad (2)$$

非類似度  $\Delta_t$  は、  $0 \leq \Delta_t \leq 1$  の範囲の値を示し、この値が大きければ大きいほど  $t$  の期間は同行している可能性が低い事を示している。

$\Delta_t$  が類似度の閾値  $\theta_\Delta$  より小さい時に両者は同行している可能性が高く、この状態が  $c$  の期間続く時に両者は同行していると推測する。各時刻での同行指数  $\delta_t$  を求め、その和から人間関係グラフの強度  $W$  を求める。また、  $\alpha_t$  が活動量の閾値  $\theta_\alpha$  より小さい時、そもそも活動が無いと定義している。

$$W = \sum_T \delta_t \quad (3)$$

$$\delta_t = \begin{cases} 1 & \bigcap_{\tau=t}^{t+c} (\alpha_\tau > \theta_\alpha \cap \Delta_\tau < \theta_\Delta) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

連絡先: 坪内 孝太, ヤフー株式会社 Yahoo! JAPAN 研究所,  
東京都港区赤坂 9-7-1 ミッドタウンタワー, 03-6864-3412,  
ktsubouc@yahoo-corp.jp

表 1: 実験の概要

用いたデバイス	Fitbit one
期間	2013.06.26 - 2013.07.09
被験者	31名
組織	企業(3部署)あるいは大学関係者
年齢	21 - 42歳

### 3. 実証実験

#### 3.1 実験の概要

実験の概要を表 1 に示す。大学関係者あるいは企業に所属しているユーザの計 31 名を対象に約 2 週間の実験を行った。31 名はそれぞれ、10 名が大学の研究室、2 名は大学の同研究室卒業生、11 名は企業の研究所、7 名は同企業の開発部門 A、1 名は同企業の開発部門 B に所属している。また大学研究室メンバの 10 名のうち 1 名が教員である。両組織は地理的に離れた場所に位置している。

万歩計には Fitbit one という機器を用いた。Fitbit one がスマートフォンからデータを逐次サーバにアップロードする機能を持った万歩計であり、データ収集が容易なためである。

また、被験者には万歩計の装着を徹底したが、その他は特に意識せず平常通りの生活を過ごしてもらうよう指示した。

#### 3.2 本手法の評価

提案手法の評価として、定量および定性的な 2 種類の評価を行う。

定量的な評価として同行検出の精度を調べる。被験者全員に期間中毎日アンケートを行い、実際同行した人をヒアリングする。ヒアリング結果と本手法により検出された同行がどの程度一致するかを調べる。具体的には、様々なパラメータセットでシミュレーションを行い、精度 (precision) および再現率 (recall) を測定し、 $w, \theta_\alpha, \theta_\Delta, c$  を変えつつ、AUC (Area Under the Curve) にて評価する。

定性的な評価は、最終的に導出される人間関係グラフについて現実の人間関係と比べ、考察する。

#### 3.3 実験の結果

まず、定量的な評価であるが、AUC は 0.72 という値を示した。パラメータの事例の一例を示すと同行検出において precision=0.85、recall=0.60 程度であった。既存研究 [Justin 10] の性能と比較すると十分な性能が出ていることがわかる。

次に、定性的な評価を行う。10 日間の実験で生成された人間関係グラフを図 2 に示す。図中の人間関係グラフは、 $w=60$ 、 $\theta_\alpha=5500$ 、 $\theta_\Delta=0.05$ 、 $c=15$  というパラメータを用いた際の結果である。各ノードが 1 ユーザを表し、ユーザ同士を結ぶリンクの太さが両者の人間関係の強さを表している。ノードが置かれている位置自体に意味はないが、関係が強い人同士が近く、力学モデルのアルゴリズムで自動的に描画している。

結果を見ると、高精度に人間関係グラフを描くことができているとわかる。

まず、各組織および各部署はうまく分かれている。企業の研究所および開発部門ではさらに細かなチームに細分化されるが、それらもうまく表現されている。さらに、企業の研究所と研究室とは共同研究を行っているが、共同研究に関係しているユーザが両組織の周辺部分に集まっている。得られる人間関係グラフとしては十分な精度といえる。

一方、僅かではあるが誤検出もみられる。たとえば、研究室 OB 生と企業の研究所との間、および大学の研究室メンバ

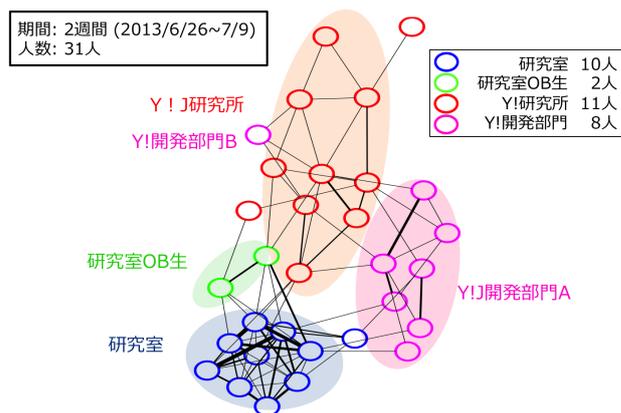


図 2: 歩数データから作成した人間関係グラフ

と企業の開発部門との間に若干リンクが生じている。しかし、両者に面識はない。推定される両者のリンクの強さ  $W$  は小さく、たまたま似た時刻に同じ行動をしたノイズによる誤検出と考えられる。

### 4. 結論

万歩計のデータから実世界の人間関係グラフを生成するアルゴリズムについて提案し、その性能を被験者実験により評価した。

提案手法は、被験者の万歩計データの各区間の歩数パタンの類似度から同行しているかどうかを検出し、検出された同行の頻度に応じて人間関係の強度を定義するという手法である。

提案した手法を 31 名の被験者に試し評価した結果、本手法の有効性を確認する事ができた。手法自体はとても単純であるが、AUC による評価で高性能なモデルが生成されていることがわかった。また、定性的な評価でも、実際の人間関係をうまく再現できていることを確認した。

今後の課題としては、本稿で提案する手法をベースにさらにロバストな手法を構築していく事があげられる。本手法で得られる人間関係はパラメータによるものが大きい。最適なパラメータが被験者数や期間といった問題規模によって異なる。問題規模や利用者の特性によらず、最適なパラメータを導出できる計算方法を考案することで本手法が現実的なものになると考えられる。

### 5. 謝辞

他の業務があるにも関わらず、実験に協力いただいた被験者の皆様に感謝します。

### 参考文献

[Justin 10] Justin Cranshaw, et. al. "Bridging the gap between physical location and online social networks", In Proc. of the 2010 ACM on Pervasive and Ubiquitous Computing, pp. 119-128, ACM, 2010.

[Jacopo 12] Jacopo Staiano, et. al. "Friends don't lie: Inferring personality traits from social network structure", In Proc. of the 2012 ACM on Pervasive and Ubiquitous Computing, pp. 321-330, ACM, 2012.