

# 同行時間の多様性を考慮した線形識別モデルに基づく 活動量計の歩数データからの同行検出法

下坂 正倫<sup>1,a)</sup> 荒木 青伊良<sup>1,b)</sup> 川尻 亮真<sup>1,c)</sup> 坪内 孝太<sup>2,d)</sup>

概要：本講演では、昼食/夕食を共にする、といった人間関係の親密度の一指標といえる同行を、ウェアラブル活動量計で得られる歩数データから頑健に検出する手法を提案する。二者の歩行を伴う同行は、二者の静的な近接に比べ、より自主的で親密度の高い交流といえる。我々はこれまで、二者間の時系列歩数データの類似性に注目することで、歩数データから同行の検出が可能であることを示してきた。しかしながら一定の時間幅に注目した類似度を利用していたため、例えば長時間にわたる懇親会と勤務中の短時間の昼食といった、異なる時間幅で生じる同行に対しては検出性能が低下する問題があった。この課題に対し本研究では、複数の時間幅に注目した二者間の歩数データの類似度の重み付けから検出を行う線形識別モデルを提案する。複数組織に属する 31 名 11 日間の行動データから同行検出性能の評価を行い、提案手法の有効性を確認した。

キーワード：人間関係推定、活動量計、歩数データ、同行検出

## 1. 序論

簡易に持ち運びできる活動量計の開発が近年盛んに行われている。それに伴い活動量計で取得できるデータを用いた研究が盛んである [1], [2]。活動量計には加速度センサや高度計が組み込まれており、歩数・カロリー消費量・階段昇数・睡眠の長さなどを装着するだけで簡単に取得することが出来る。Fan らは活動量計 Fitbit<sup>\*1</sup> で取得できるデータをアートやグラフで表現することで、運動の楽しさや運動量を増やすことが出来るとした [2]。Montgomery-Downs らは、腕の動きの加速度を活動量計 Fitbit で取得し睡眠状態を推定するとき、脳波等を用いた睡眠ポリグラフ検査によって睡眠状態を計測するときの結果の比較を行った [3]。このように活動量計のアプリケーションは個々人の健康支援に主眼がおかれている。一方 Tsubouchi ら [4]・川尻ら [5] は活動量計の新たなアプリケーションとして、昼食/夕食を共にするといった人間関係の親密度の一指標といえる同行の、長さや頻度に基づく人間関係推定手法を

提案した。活動量計の 1 分毎の歩数の時系列データ（歩数データ）の二者間の類似性に注目することで、歩数データから同行の検出が可能である。

ビジネスの現場で face-to-face の関係が重要であること [6] から、従来も様々なデータを用いた face-to-face の人間関係推定手法が提案されていた。これらは二者が同じ時間に同じ場所にいることを計測し関係を推定しており、GPS・Wi-Fi を用いた手法 [7], [8] や Bluetooth を用いた手法 [9], [10]、赤外線を用いた手法 [11] があった。活動量計の歩数データを用いた人間関係推定手法 [4], [5] でも face-to-face の交流を推定する。活動量計の歩数データを用いるため、地図上の位置が特定される GPS・Wi-Fi を用いる手法に比べると比較的プライバシー低侵害で、電力消費も小さく済み、Bluetooth・赤外線を用いるときのように特殊な機器や設定を必要としないといったメリットがある。

本研究では既存の歩数データを用いた人間関係推定 [4] の性能向上を念頭に、長時間にわたる懇親会と勤務中の比較的短い昼食といった、異なる時間幅の同行について検出性能の頑健化に取り組む。先行研究 [4] では同行の検出において、一定の時間幅に注目した二者間の歩数データの類似度のみを利用していたため、このような異なる時間幅の同行に対しては検出性能が低下する問題があった。本研究ではこの課題に対処し同行の時間幅の多様性を考慮した検出手法を構築する。

<sup>1</sup> 東京大学  
Tokyo University

<sup>2</sup> Yahoo! JAPAN 研究所  
Yahoo! JAPAN Research

a) simosaka@ics.t.u-tokyo.ac.jp

b) araki@ics.t.u-tokyo.ac.jp

c) kawa.jiri@ics.t.u-tokyo.ac.jp

d) ktsubouc@yahoo-corp.jp

\*1 <http://www.fitbit.com/one>

## 2. 同行の時間幅の多様性を考慮した検出手法

本章では同行の時間幅の多様性を考慮した検出手法として、複数の時間幅に注目して得た複数の二者間の歩数データの類似度を、最適な重み付けで加算したスコアによって同行検出する線形識別モデルを提案する。

### 2.1 歩数データを用いた同行検出の問題設定

本研究では同行検出を、1日にある二者が同行を行ったかどうかの検出とする。これは、細かい時間単位での教師データ・評価データの収集ができないからである。人は通常、何時何分から何時何分まで誰と同行したかを、詳しく覚えておくことはできない。しかし1日にある人と同行したかどうかは回答することができる。

ある1日における二者の組み合わせを  $i = 1, \dots, n$  で表す。 $n$  は二者の組み合わせパターンと取りうる日数の積となる。つまり同じ二者でも異なる日には  $i$  の値は別のものとなる。この1日の二者  $i$  の歩数データを  $a_i, b_i$  と表す。これらは、1日分の1分値の歩数が並んだ時系列データのベクトルを表しており、 $a_i, b_i \in \mathbb{R}^{1440}$  である。また、二者の歩数をまとめて  $x_i$  で表し、 $x_i = \{a_i, b_i\}$  となる。時刻  $t$  での歩数値は  $a_t, b_t$  と表す。1日単位での二者  $i$  の同行の有無は、 $y_i = \{-1, 1\}$  で表す。 $-1$  のとき二者は同行していないとし、 $1$  のとき二者は同行したとする。この  $y_i = \{-1, 1\}$  が取得できる教師データであり、本研究で推定する対象である。

### 2.2 同行時間の多様性を考慮した同行検出の定式化

同行の時間幅の多様性を考慮した検出手法として、複数の時間幅に注目して得た複数の二者間の歩数データの類似度を最適な重み付けで加算したスコアを用いて、1日単位での同行検出を行う方法が考えられる。例えば60分の時間幅に着目した類似度のみを用いた場合、類似度の値は60分の時間幅の同行データのみで大きくなるため、60分の時間幅の同行しか検出できない。しかし60分の時間幅に着目した類似度と、120分の時間幅に着目した類似度を加算したスコアを用いれば、60分・120分の時間幅の同行データどちらともにおいてスコアの値が大きくなるので、どちらの時間幅の同行も同時に検出できるということである。このとき、注目する時間幅の同行を表す類似度を算出するための最適なパラメータは時間幅毎に異なるため、注目する時間幅毎に類似度を最適化する必要がある。先行研究 [4] の枠組みをそのまま利用しようとする、同行の時間幅毎に類似度を最適化するために同行の時間幅に関する教師データが必要であるが、2.1 で述べたように同行した時間を細かく覚えておくことは出来ず同行した時間幅の情報は得られない。一方で線形識別モデルを用いれば、複数の時間幅に着目したとき各時間幅毎に最適化した類似度を

自動的に選択し、各時間幅毎の最適な類似度を加算したスコアを自動的に得ることができる。必要な教師データは1日単位での同行の有無のみで良い。

線形識別モデルを用いた同行検出は、ある1日での二者  $i$  についての同行のスコア  $f(x_i) = w^T \phi(x_i)$  を用いて、1日単位での二者  $i$  の同行の有無を検出する。1日の二者  $i$  のデータ  $x_i$  から、様々な時間幅で算出した類似度の複数の値を並べて、特徴ベクトル  $\phi(\cdot)$  をつくる。このとき注目する時間幅毎の最適な類似度は不明のままでよく、時間幅毎に複数のパラメータで算出した類似度を特徴ベクトルに並べる。この特徴ベクトル  $\phi(\cdot)$  と重み  $w$  との内積をとることで、時間幅毎の最適な類似度に正の重み付けがされ、様々な時間幅での最適な類似度を加算したスコアを得ることができる。重み  $w$  と特徴ベクトル  $\phi(\cdot)$  の内積が同行を判定する閾値より大きければ二者  $i$  はその1日に同行したと判定し、以下であれば二者  $i$  はその1日に同行していないと判定する。

### 2.3 同行検出性能の最適化

時間幅毎に最適な類似度を選択し複数の時間幅の類似度を加算するための重み  $w$  は  $w = \operatorname{argmin} \sum_i l(f(x_i), y_i) + \lambda R(w)$  を用いて、推定誤差の関数  $l(f(x_i), y_i)$  を最小化するように求める。 $\lambda \geq 0$  は正則化係数、 $R(w)$  は正則化項である。本研究では、推定誤差の関数としてヒンジ損失関数 [12] を用い、正則化項として、疎な重みを得られる L1 正則化を用いた。 $R(w) = \sum_i |w_i|$  となる。また最適化のアルゴリズムとして、疎な重みを得られ実装の容易な FOBOS [13] を用いた。提案手法における FOBOS のアルゴリズムは以下ようになる。ここで、 $k$  は劣勾配の計算に用いるサンプルの数、 $\lambda$  は正則化係数、 $T$  は重みの更新回数である。 $|\cdot|_+$  はヒンジ損失関数を指す。 $\operatorname{sign}(\cdot)$  は、引数の正負の値を返す。 $w_t$  の  $l$  次元目の要素を  $w_{t,l}$  で表す。

---

#### Algorithm 1 FOBOS

---

```

INPUT  $S = \{x_i, y_i\}_{i=1}^n, \lambda, T, k$ 
OUTPUT  $w_{T+1}$ 
initialize Choose  $w_1$  s.t.  $\|w_1\| \leq 1/\sqrt{\lambda}, |w| = |x_i|$ 
for  $t = 1, 2, \dots, T$  do
  Choose  $A_t \subseteq S$ , where  $|A_t| = k$ 
  Set  $A_{(+t)} = \{(x_i, y_i) \in A_t : y_i w_t^T x_i < 1\}$ 
  Set  $\eta_t = 1/\lambda t$ 
  Set  $w_{t+1/2} = w_t + \eta_t \sum_{(x_i, y_i) \in A_{(+t)}} x_i y_i / k$ 
  Set  $\eta_{t+1/2} = 1/\lambda t$ 
  for  $l = 1, 2, \dots, |w_{t+1/2}|$  do
    Set  $w_{t+1,l} = \operatorname{sign}(w_{t+1/2,l}) |w_{t+1/2,l} - \eta_{t+1/2} \lambda|_+$ 
  end for
end for

```

---

### 3. 二者間の歩数データの類似度の設計

提案手法では同行の特徴を示す複数種類の類似度を利用でき、類似度として次の4種類を用いる。

- 歩数値一致度の和
- 歩行タイミング一致度の和
- 速度一致度の和
- 活動レベル

1つ目の歩数値一致度の和は先行研究 [4] で同行検出に利用されたものである。残りの3種類は本研究で新たに設計したものである。

#### 3.1 歩数値一致度の和

歩数値一致度の和は先行研究 [4] で同行検出に利用された二者間の歩数データの類似度である。同行時に同じ時刻における二者の歩数値が近い値をとるに基づいており、一定の時間幅における二者の歩数差の小ささを表している。提案手法では複数の時間幅に対して複数の歩数値一致度の和を算出し、それぞれ特徴ベクトルに並べる。

一定の時間幅  $w$  の二者の歩数データに着目する。まず、時刻  $t$  における活動量  $\alpha(a, b)_t$  を式 (1) で定義する。活動量が閾値  $\theta_\delta$  より小さいとき、睡眠中やデスクワーク中とみなし、二者は同行していないとする。

$$\alpha(a, b)_t = \sum_{\tau=t}^{t+w} a_\tau^2 + b_\tau^2 \quad (1)$$

そうでないとき、時刻  $t$  の差異値  $\Delta(a, b)_t$  は式 (2) で表される。

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

これらを用いて、時刻  $t$  における歩数値一致度  $\delta_t$  は式 (3) で計算される。 $\theta_\delta$  は二者間の歩数データの差異値  $\Delta_t$  の閾値であり、差異値が大きいとき、二者は別の行動をしていると考えられる。さらに継続時間のパラメータ  $c$  を用い、以上の同行の条件を満たす時刻がどれだけ継続するかによって、一定の時間幅  $w$  における時刻  $t$  での歩数値一致度  $\delta_t$  が求まり、1日分加算して歩数値一致度の和となる。

$$\delta_{t+w/2} = \begin{cases} 1 & (\forall \tau (\tau = t, \dots, t+c \rightarrow \alpha_t > \theta_\alpha \cap \Delta_t < \theta_\delta)) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

#### 3.2 歩行タイミング一致度の和

歩行タイミング一致度の和は、歩行開始の一致度・歩行停止の一致度・歩行終了の一致度の3種類からなり、以下でそれぞれの算出方法を説明する。

歩行開始の一致度は式 (4) で表される。歩行開始の一致度は、二者が共に歩いていない状態から、同時に歩いている状態に遷移する時刻に着目している。一緒にデスクワー

クをしていて、誘い合って昼食に行くなどの行動を表すことができる。ここで  $\tau_{ab}$  は取りうる時刻の最小値である。図1は歩行開始の一致度を示しているが、二者の歩行開始のタイミングが一致したときに値が大きくなっていることがわかる。

$$s_t = \begin{cases} t - \tau_{ab} & ((\forall \tau (\tau = \tau_{ab}, \dots, t-1 \rightarrow a_\tau, b_\tau \leq \theta)), \\ & (a_t > \theta), (b_t > \theta)) \\ 0 & (\text{otherwise}) \end{cases} \quad (4)$$

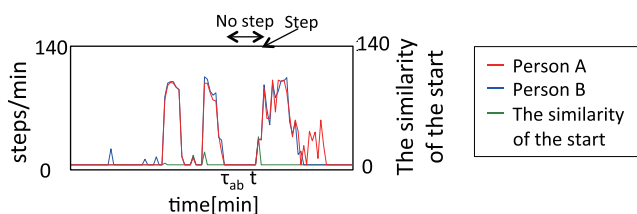


図1 歩行開始の一致度

次に、歩行停止の一致度は式 (5) で表される。歩行停止の一致度は、二者が共に歩いている状態から、同時に歩いていない状態に遷移する時刻に着目している。一緒に店へ歩いていて、到着して着座したなどの行動を表すことができる。ここで  $\tau_{ab}$  は取りうる時刻の最小値である。

$$o_t = \begin{cases} t - \tau_{ab} & ((\forall \tau (\tau = \tau_{ab}, \dots, t-1 \rightarrow a_\tau, b_\tau > \theta)), \\ & (a_t \leq \theta), (b_t \leq \theta)) \\ 0 & (\text{otherwise}) \end{cases} \quad (5)$$

最後に、歩行終了の一致度は式 (6) で表される。歩行終了の一致度は、二者が共に歩いている状態から同時に歩いていない状態に遷移する時刻に着目し、二者ともに歩いていない状態がどれほど続くかによって同行の終了を表している。一緒に店から職場に戻って来て、再びデスクワークに戻るといった行動を表すことができる。

$$e_t = \begin{cases} \tau_a - t & ((\forall \tau (\tau = t+1, \dots, \tau_{ab} \rightarrow a_\tau, b_\tau \leq \theta)), \\ & (a_t > \theta), (b_t > \theta)) \\ 0 & (\text{otherwise}) \end{cases} \quad (6)$$

以上のようにして求めた歩行タイミングの一致度を、それぞれ1日分加算する。これをそれぞれ特徴ベクトルに並べる。

#### 3.3 速度一致度の和

速度一致度の和は、同行時は二者の歩行速度が類似するという考えに基づく。歩数値一致度の和では二者の歩数差が時間幅  $w$  で平均化されるので、図 (2)・図 (3) に示すような速度が異なるときも同じときも、ほぼ同じ値をとる可

能性がある．このようなパターンを区別するために重要であると思われる．

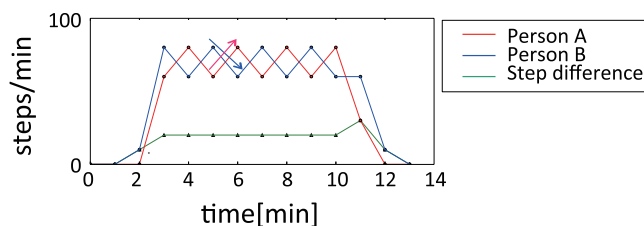


図 2 同行していない二者の歩数データ

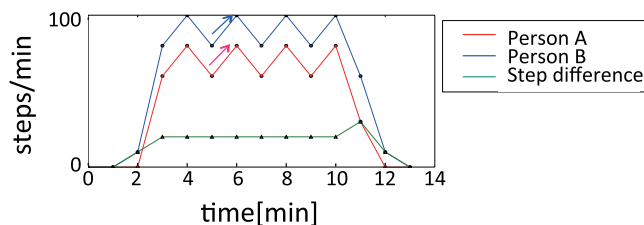


図 3 同行している二者の歩数データ

時刻  $t$  の速度類似度  $d_t$  は式 (9) で表される．

$$v_a = \frac{d}{dt} \mathbf{a} = a_{t+1} - a_t \quad (7)$$

$$v_b = \frac{d}{dt} \mathbf{b} = b_{t+1} - b_t \quad (8)$$

$$d_t = (1 - |v_a - v_b| / |v_a + v_b|) \cdot \text{sign}(v_a) \cdot \text{sign}(v_b) \quad (9)$$

この値を 1 日分加算することで，速度一致度の和を求める．

### 3.4 活動レベル

活動レベルは，同行時は 1 人で歩くときよりも比較的ゆっくり歩くという考えに基づき，どの範囲に歩数値が属するかを示すものである．歩数値の範囲，1 分間に 0 から 150 歩程度を  $H$  個の段階に分ける．二者が  $h$  段階目の歩数をとる時刻の 1 日分の和を，活動レベル  $\phi_h$  と定義し，式 (10) で示す．1 ~  $H$  段階のそれぞれでの活動レベル  $\phi_1 \sim \phi_H$  を算出し，特徴ベクトルにならべて用いる．

$$\phi_h = \begin{cases} +1/ & ((150 \cdot (h-2)/H \leq a_\tau \\ & < 150 \cdot (h-1)/H) \cup \\ & (150 \cdot (h-2)/H \leq b_\tau \\ & < 150 \cdot (h-1)/H), \\ & \tau = 1, 2, \dots, 1440) \\ +0 & (\text{otherwise}) \end{cases} \quad (10)$$

## 4. 歩数データによる同行検出性能の評価実験

本実験では提案手法によって同行時間の多様性を考慮した同行検出が出来ていることを示す．そのために一定の時間幅に注目し同行検出を行う先行研究手法 [4] と，複数の時間幅を考慮し同行検出を行う提案手法について，時間幅が異なる同行毎の検出性能を比較する．先行研究手法 [4] では注目している時間幅と異なる時間幅の同行の検出性能が低くなるが，提案手法では複数の時間幅を考慮することで異なる時間幅の同行を同時に検出できることを示す．

### 4.1 収集データ

本研究では同行検出性能を評価するため，歩数データと正解データを収集した．ここで 2 章で述べたように，同行検出とは，正解データの得られる 1 日単位での二者間の同行の有無の検出のことである．表 1 にその概要を示す．歩数データは，活動量計 Fitbit を用いて収集した．参加者がデバイスを装着するだけでデータがサーバーに自動的にアップロードされるため，解析者がサーバーから 1 分毎の歩数を取得した．

正解データは，参加者全員にアンケートを行うことで取得した．具体的には毎日夕方に，その日同行した相手をすべて回答してもらい，1 日に同行のあるペアの情報を取得した．さらに同行と回答しなかった相手をその日に同行のない相手とすることで，1 日に同行のないペアの情報を取得した．アンケートには回答に参加者同士で相違があり正解データの正確さを期すため，二者間の回答が一致しているデータのみ用いる．その場合，1 日の同行の有無に関して，同行があるデータは 282 個，同行が無いデータは 4481 個である．1 つのデータは，1 日の二者の歩数データを指す．

表 1 収集データ概要

Period	Weekdays from 2013/6/26 to 2013/7/10
Member	31
The number of organization	5
Age range	21yrs to 42yrs

また本実験では時間幅が異なる同行毎の検出性能を比較するが，同行の時間幅を直接尋ねることは回答者の記憶力の点で難しかったため，今回は参加者に同行の目的を回答してもらい目的の異なる同行毎に検出性能を比較する．同じ同行の目的であれば同行の時間幅は似たものになる．例えば，昼食の目的であれば昼休み中に終わるために 1 時間弱程度の時間幅になり，夕食・懇親会の目的であれば昼食より長めの時間幅になり，喫煙の目的であれば 30 分弱程度の短めの時間幅になる．同行の時間幅でなく同行目的を尋ねることに，回答の手間・回答者の心理的負担という問題点があるが，実際に提案手法を運用する際は必要ないも

のである．具体的に参加者に尋ねる同行の目的は，昼食・夕食・喫煙とした．1日に複数の目的の同行をしたデータは除外し，単一の目的の同行を行ったデータのみを用いる．データ数は同行の目的毎に，昼食：134，夕食：53，喫煙：7である．

#### 4.2 同行時間の多様性を考慮した同行検出実験方法

この実験では提案手法によって同行時間の多様性を考慮した同行検出が出来ていることを示す．実験の流れは以下の通りである．まず，先行研究 [4] 手法において注目する時間幅毎（今回は近似的に目的毎）に最適な類似度を求め，そのときの目的毎の同行有データの最大の同行検出性能を算出する．またこのとき注目した目的と異なる目的の同行有データの同行検出性能は低下していることを確認する．次に提案手法においては，同行目的の情報は用いず1日の同行検出性能が最大となるような重み  $w$  を求め，その重み  $w$  を用いて目的毎の同行有データの同行検出性能を算出する．先行研究手法では注目する時間幅（目的）と同じ目的の同行有データの同行検出性能が最大となるが，提案手法では複数の時間幅を考慮できていれば，先行研究手法での最大同行検出性能を同時に示せるはずである．

目的毎の同行有データの同行検出性能は，4.1 で収集した全データの同行検出性能の precision が 0.8 以上になるとき，目的毎の同行有データを同行と検出する率である recall の最大値を求めて評価する．同行判定の閾値を厳しくするとこの precision は高くなり recall は低くなるというように，これらはトレードオフの関係になっている．recall 値が 1 に近いほど，同行検出性能が良い．

先行研究 [4] では具体的には以下のように同行検出を行った．まず一定の時間幅に注目した二者間の歩数データの類似度である歩数値一致度の和を求め，注目する時間幅毎（今回は近似的に目的毎）に最適な類似度，歩数値一致度の和を求めパラメータを選択するため，収集したデータを日付で分け 5 交差検定を行う．時間幅  $w$  ・非類似度の閾値  $\theta_\delta$  ・継続時間の閾値  $c$  といったパラメータは相互に関係し合っており，またデータから推測することは出来ないで，このような最適化が必要である．訓練データにおいて目的毎の同行有データの同行検出性能が最大となるように，表 2 から時間幅  $w$  などのパラメータ選択を行う．活動量の閾値  $\theta_\alpha$  は，先行研究で用いた 5500 を利用した．求めた歩数値一致度の和がある閾値より大きいならばその日は同行，そうでないならばその日は不同行とする．

提案手法においても 5 交差検定を行い，1日の同行検出性能が最大になるように，様々な時間幅に着目して得られる類似度の複数の値に最適に重み付けするための重み  $w$  を求める．提案手法では新しい種類の類似度の値を加えて利用できるため，3章で定義した以下の類似度を用いる．括弧内は省略記述である．表 3 はそれらの類似度を算出する

表 2 歩数値一致度の算出に用いるパラメータの選択肢

時間幅 $w$	5, 10, 30, 60, 90, 120, 180
非類似度の閾値 $\theta_\delta$	0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4
継続時間の閾値 $c$	1, 5, 15

際に用いたパラメータと，特徴ベクトルに並べる数を表している．

- 歩数値一致度の和 (original)
- 歩行タイミング一致度の和 (timing)
- 速度一致度の和 (velocity)
- 活動レベル (activity)

表 3 用いた二者間の歩数データの類似度

二者間の歩数データの類似度	パラメータ		特徴ベクトルに並べる数
original	時間幅 $w$	5, 10, 30, 60, 90, 120, 180	210
	非類似度の閾値 $\theta_\delta$	0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4	
	継続時間の閾値 $c$	1, 5, 15	
timing	歩数の閾値 $\theta$	0, 10, 20	9
velocity	-	-	1
activity	歩数値の範囲の数 $H$	7	7

#### 4.3 同行時間の多様性を考慮した同行検出実験結果

表 4 は提案手法と先行研究手法 [4] での目的毎の同行有データの同行検出性能を示している．先行研究手法では，注目する時間幅毎（今回は近似的に目的毎）に類似度を最適化したとき，その目的の同行有データを同行と検出する率が増え recall 値が大きくなっている．しかしそれ以外の目的の同行有データの recall 値は低下している．例えば長めの時間幅を持つと予想される夕食に注目したとき，夕食の同行有データの recall 値が 0.62 と最大になっている．相対的に短めの時間幅である昼食の同行有データの recall 値が最大値 0.65 から小さくなって 0.63 になっている．さらに短い時間幅である喫煙の同行有データの recall 値が最大値 0.57 から小さくなって 0.14 になっている．

ここで提案手法を見ると，昼食・夕食・喫煙の同行有データの recall 値が同時に 0.82, 0.66, 0.29 となっている．昼食・夕食の同行有データの recall 値については先行研究手法の最大値 0.65, 0.62 より良い性能を同時に示せている．図 4 からこのことを示している．一方，提案手法において喫煙の同行有データの recall 値は，先行研究手法の最大値 0.57 より小さく 0.29 である．しかし先行研究手法で夕

食の同行有データの recall 値が最大 0.62 のとき喫煙の同行有データの recall 値は 0.14 であり，提案手法では夕食の同行有データの recall 値が 0.66 のとき喫煙の同行有データは 0.29 なので，先行研究手法 [4] に比べると時間幅の異なる同行を同時に検出できている．以上から提案手法によって同行時間の多様性を考慮できていることが示された．

表 4 同行目的毎（同行時間幅毎）の 1 日単位で同行検出性能

評価用のデータ	先行研究手法 [4]			提案手法
	注目した目的（時間幅） （用いた時間幅 $w$ の平均）			
	喫煙 (30 分)	昼食 (72 分)	夕食 (132 分)	様々な時間幅を考慮
昼食	0.49	<b>0.65</b>	0.63	<b>0.82</b>
夕食	0.45	0.60	<b>0.62</b>	<b>0.66</b>
喫煙	<b>0.57</b>	0.29	0.14	0.29

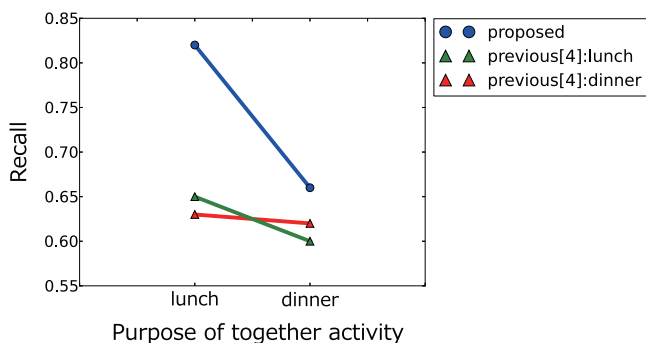


図 4 提案手法と先行研究手法 [4] での目的毎の同行検出性能

## 5. 結論

本研究では既存の歩数データを用いた人間関係推定 [4] の性能向上を念頭に，同行検出性能の頑健化に取り組んだ．先行研究 [4] では異なる時間幅の同行に対しては検出性能が低下する問題があったため，同行の時間幅の多様性を考慮した検出手法として識別モデルを提案した．

複数組織に属する 31 名 11 日間の活動量計の歩数データから時間幅毎の 1 日単位での同行検出性能の評価を行った．先行研究手法 [4] においては，注目する時間幅（今回は近似的に目的）毎に同じ目的の同行有データの検出性能が最大となり，昼食・夕食・喫煙の同行有データの recall 値がそれぞれ最大で 0.65・0.62・0.57 となった．一方提案手法においては，昼食・夕食・喫煙の同行有データの recall 値が同時に 0.82・0.66・0.29 となった．昼食・夕食については提案手法の方が先行研究手法より良い性能を同時に示している．喫煙については先行研究手法の最大値には及ばないが，先行研究手法で昼食・夕食に注目したとき喫煙の同行有データの recall 値は 0.29・0.14 であり，提案手法ではこれら以上の性能で検出できている．以上から先行研究手法で検出性能が低下する異なる時間幅の同行を，提案手法によって同時に検出できたと考えられ，複数の時間幅に

着目する提案手法の有効性を示した．

今後の課題としては関係推定に重要な同行した時間帯を推定することが挙げられる．しかし時間単位での同行に關する教師データは収集できないため，最適化方法に工夫が必要である．

## 参考文献

- [1] Jupe, N., Faries, M., Jones, E. and Whitehead, M.: Evaluation of the Nike+ FuelBand in Energy Expenditure and Steps Taken during Exercise, *In Proc. of IJESAB 2014*, p. 71.
- [2] Fan, C., Forlizzi, J. and Dey, A.: A Spark of Activity: Exploring Informative Art as Visualization for Physical Activity, *In Proc. of Ubicomp 2012*, pp. 81–84.
- [3] Montgomery-Downs, H., Insana, S. and Bond, J.: Movement toward a Novel Activity Monitoring Device, *Sleep and Breathing*, Vol. 16, No. 3, pp. 913–917.
- [4] Tsubouchi, K., Kawajiri, R. and Shimosaka, M.: Working-Relationship Detection From Fitbit Sensor Data, *In Adjunct Proc. of UbiComp 2013*, pp. 115–118.
- [5] 川尻亮真, 坪内孝太, 荒木青伊良, 下坂正倫: 活動量計データと用いた同行検出手法の研究, 人工知能学会 2014.
- [6] Kirkman, B., Rosen, B., Tesluk, P. and Gibson, C.: The Impact of Team Empowerment on Virtual Team Performance: the Moderating Role of Face-to-face Interaction, *Academy of Management Journal*, Vol. 47, No. 2, pp. 175–192.
- [7] Crandall, D., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D. and Kleinberg, J.: Inferring Social Ties from Geographic Coincidences, *In Proc. of NAS 2010*, pp. 22436–22441.
- [8] Pham, H., Shahabi, C. and Liu, Y.: EBM - An Entropy-Based Model to Infer Social Strength from Spatiotemporal Data, *In Proc. of SIGMOD 2013*, pp. 265–276.
- [9] Madan, A., Cebrian, M., Lazer, D. and Pentland, A.: Social Sensing for Epidemiological Behavior Change, *In Proc. of UbiComp 2010*, pp. 291–300.
- [10] Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N. and Pentland, A.: Friends Don't Lie: Inferring Personality Traits from Social Network Structure, *In Proc. of UbiComp 2012*, pp. 321–330.
- [11] Choudhury, T. and Pentland, A.: The Sociometer: A Wearable Device for Understanding Human Networks, *In Proc. of CSCW 2002*.
- [12] Gentile, C. and Warmuth, M.: Linear Hinge Loss and Average Margin, *In Adv. of NIPS 1998*, pp. 225–231.
- [13] Duchi, J. and Singer, Y.: Efficient Online and Batch Learning Using Forward Backward Splitting, *Journal of Machine Learning Research*, Vol. 10, pp. 2899–2934.