

日常生活習慣パターン把握に基づく スポーツジムの新たな指導サービスの創出

Recommendation Service based on Lifestyle Patterns extracted from User Profile Data

荒木 次郎*¹ 小池 亜弥*² 松崎 和賢*² 小関 悠*² 白井 康之*²
 Jiro Araki Aya Koike Kazutaka Matsuzaki Yu Koseki Yasuyuki Shirai

*¹(株) 三菱総合研究所 先端科学イノベーショングループ
 Mitsubishi Research Inst. Inc. Advanced Science and Innovation Group
 *²(株) 三菱総合研究所 先端情報技術研究グループ
 Mitsubishi Research Inst. Inc. Advanced Information Technology Group

In order to create new recommendation service in sports club, we propose a method to extract lifestyle patterns by integrating health and positional information by several sensors, and records of food and exercises. In this paper we show results of attribution analysis for behavior-change motivation and recommendation method based on lifestyle patterns.

1. はじめに

スポーツジムは過渡期を迎えている。これまでのスポーツジムのサービスはプールやジムマシン、プログラムの提供など、基本的に場の提供が主であった。しかし趣味の多様化やジム業界の飽和により、これまでのサービスだけでは続けられなくなっている。そのため、各企業はさまざまな試みを行っており、プログラムの充実や施設のラグジュアリー化など場の充実とともに、運動指導や身体変化の記録サービスなどのソフトウェア面などの工夫も試みられている。

本研究では、スポーツジムの指導サービスを情報科学によって支援することを目的としている。本稿では、情報大航海プロジェクトの中で開発したプロフィール情報解析基盤の実証実験として、スポーツジムにおけるプロフィール情報収集と、プロフィールから推定されるユーザ特性を考慮した運動指導サービスの将来像について検討する。

2. 実証実験の概要

本章では、経済産業省の「情報大航海プロジェクト」の中で我々が行った「プロフィール情報解析基盤の構築」とその検証実験について紹介する。

2.1 プロフィール情報解析基盤 IGVminer

情報大航海プロジェクトは、Web 上のテキスト情報に限らず、映像や音声のようなリッチコンテンツ、そしてユーザの実世界の行動履歴や嗜好、健康情報などを扱う検索・解析技術を開発することを目指したプロジェクトである。汎用的な共通基盤を構築する取り組みと、このような基盤を用いた次世代サービスの可能性を検証する実証事業などから成る。我々は、NTT アドバンステクノロジー、有限責任中間法人 Pucc、ビジネスマイニング研究センターらと共同で、共通基盤構築の1つの課題として、個人の属性や行動の履歴（プロフィール）を収集・蓄積・解析するための基盤を構築してきた。基盤は、プロフィール情報収集基盤、プロフィール情報統合基盤、プロフィール情報解析基盤から構成される。

連絡先: 荒木 次郎, (株) 三菱総合研究所 先端科学イノベーショングループ, 〒 100-8141 東京都千代田区大手町 2-3-6, jiro@mri.co.jp

このうちプロフィール情報解析基盤では頻出パターンマイニング実行環境 IGVminer を開発し、以下のような頻出パターンマイニングのアルゴリズムを実装している。

- LCM (Linear time Closed itemset Miner)
 宇野らによって開発された、非常に高速な頻出アイテム抽出アルゴリズム [Uno03]
- VSOP (Value-Sum-of-Products Calculator)
 湊らによって開発された、抽出パターンの高速な比較・検索アルゴリズム [湊 05]

また、一般の方がより使いやすくサービス寄りの利用が想定されるクラス分類ツールを開発した。これらはいずれも IGVminer の頻出パターン抽出をベースとしてクラス分類器を生成することができる。

- IGVcaep: 顕在パターンの集約スコア判定法
 LCM の高速頻出パターン抽出機能を利用して、CAEP (Classification by Aggregating Emerging Patterns) [Dong99] に基づく分類を行う。CAEP は、各クラス (C,C') の顕在パターン EP(Emerging Pattern) の集約スコア (Aggregate Score) を基に、クラス判定を行う方法である。式 (1)~(3) から分かるように、本手法はクラスに特徴的なパターンを複数もつことによる重み付き多数決判定法である。なお、式 (2) に示すように集約スコアは学習データ数やパターン数の偏りの影響をなくすために学習データのスコア中央値で正規化している。

データ s のクラス C 集約スコア:

$$score(s, C) = \sum_{e \subseteq s, e \in EP(C)} \frac{GR_C(e)}{GR_C(e) + 1} \times SP_C(e) \quad (1)$$

$$SP_C(e) = \text{パターン } e \text{ を含むクラス } C \text{ のデータ比率 (支持度)}$$

$$GR_C(e) = \frac{SP_C(e)}{SP_{C'}(e)} \quad (\text{パターン } e \text{ のクラス } C \text{ 増加率})$$

データ s のクラス C 集約スコア (正規化後):

$$\begin{aligned} norm_score(s, C) &= \frac{score(s, C)}{base_score(C)} & (2) \\ base_score(C) &= \text{クラス } C \text{ の学習データの} \\ &\quad \text{集約スコア中央値} \end{aligned}$$

データ s のクラス C 判定基準:

$$norm_score(s, C) > norm_score(s, C') \quad (3)$$

- IGVdtep: 頻出パターンを条件とする決定木判定法
LCM の高速頻出パターン抽出機能を利用して、クラス間のコントラストパターンを抽出し、コントラストパターンを条件ノードとする決定木による分類を行う。コントラストパターンは、式 (4) に示すようにクラス間の支持率の差が一定以上あるものと定義される。決定木の作成は χ^2 値に基づく方法を探っている。

$$|SP_C(e) - SP_{C'}(e)| > \text{最小支持率} \quad (4)$$

2.2 スポーツジムにおけるプロフィール情報収集実験

我々は、情報大航海プロジェクトの中で構築したプロフィール情報の収集・蓄積・解析基盤が実環境、実データに対して適用可能であることを検証するために、スポーツジムの会員サービスを想定した検証実験を行った。実験では、(株) 東急スポーツオアシスの協力を得て、モニター対象者からジム内外での運動履歴や食事内容、体重・血圧等のヘルスデータ等を収集・蓄積した。そして収集したプロフィール情報をもとに個人の嗜好、目的意識、日常生活習慣などに基づく新たな運動指導サービスの可能性を検証した。本節では、スポーツジムにおけるプロフィール情報の収集実験の概要を紹介する。

- 目的
プロフィール情報解析基盤の実環境・実データ動作検証が主目的であるが、プロフィール情報を活用したスポーツジムの運動指導サービスの将来像を描くことも目的とした。そのため、収集機器のハードウェア面は今後さまざまな進歩が期待できることも想定して、今回はユーザの計測・入力の手間や抵抗感は一先ず無視して、将来的に利用可能性のあるプロフィール情報を可能な限り多く収集することにした。
また、運動指導や食事指導は、身体計測や具体的な運動内容や食事内容とともに、心理的な面も大きく影響することから、実験期間の前後や実験期間中の意識状態もモニターに入力してもらい、身体・行動・意識の関係を分析することにした。
- モニター対象者
運動による生活改善の必要を感じている 30~40 代の男女各 15 名 (合計 30 名)
- 実験期間
平成 21 年 1 月 15~28 日 (14 日間)
- 期間中のモニター義務
 - 毎日、下記のプロフィール情報を PHS 経由で送信する
 - 事前・事後アンケートに答える
 - 期間中、可能な限り 4 回以上最寄りの東急スポーツオアシスの店舗で運動をする (なお、実験期間中及び終了後 1 ヶ月間利用可能な 10 回分の店舗利用回数券を提供)

- 収集したプロフィール情報
 - 事前アンケート
モニタリング期間前の身体情報 (体重、身長など) や、食生活、運動状況、日常の活動レベル、生活改善の意識レベルなどをアンケート形式で収集した。
 - モニタリングデータ
モニターにはヘルスセンサーと PHS を提供し、日々の身体状態を計測してもらい、食事内容、運動内容、日常生活、気分などの入力内容とともに、毎日 PHS 経由で送信してもらった。身体情報としては、体重・体脂肪率・筋肉率、血圧、心拍数をヘルスセンサーで計測し PHS 経由で収集した。食事内容としては間食を含む毎食のメニューや量を PHS 上で入力してもらい収集した。運動内容はジムでの運動種目や時間・回数、日常生活内での運動の種類や時間を PHS で入力・送信してもらった。また、取り組み度合いや気分なども PHS 上で入力してもらった。その他、睡眠・入浴時などを除いて体動計を身に付けてもらい、1 日の合計歩数や数分おきの活動量変化を収集した。さらに PHS の位置計測サービスを利用して、一日の移動距離も収集した。
 - 事後アンケート
モニタリング期間終了後に、意識や運動活動、食生活の変化、モニタリング実験への印象、期待する指導サービスなどをアンケート形式で収集した。意識や運動、食生活の変化を事後にアンケートに記入してもらう理由は、これらのデータを正解例として、モニタリングデータのみからモニター特性を推定した結果を検証できるようにするためである。
- モニタリング情報のフィードバック
 - 簡易レコメンド情報
1 週間分の情報をもとに事前に設定しておいた簡単な判定ルールとレコメンドテンプレートを使って、モニターへ体重変化、食事量、運動頻度のアドバイスをメール送信した。
 - セルフモニタリング情報
モニターから送信された日々の体重や運動量、食事量などの変化をグラフ化してモニターにメール送信した。
- 実験結果の概要
 - 集まったモニターの特徴
 - ・ 体重や食事の摂取カロリーをきちんと把握できていない人が意外に多い
 - ・ 食事は満腹まで食べ、間食をとる人が多い
 - ・ デスクワークが中心で、一日に歩く時間は 1 時間以内が大半。休日にスポーツ等で体を動かす人も多いが、ゆっくり休んでいる人もいる
 - ・ 減量を目的とする人が多い
 - モニターの実験期間中の行動特性
 - ・ 毎日きちんと計測をする人とあまりしない人に分かれる
 - ・ 運動に関してもやる人とやらない人に大きく分かれる
 - 期間中のモニターの行動・意識変化
 - ・ この短期間に体重等に有意な変化が現れた人はいないが、健康増進に役立ったと回答した人は非常に多い
 - ・ 大半の人に意識や行動に変化があった
 - ・ 半数の人は開始時に比べて改善意欲がさらに高まった
 - ・ 意識の変化としては、運動や食事、日常生活、身体への関心が高まったこと

・変化のきっかけは、実際に運動したことや、身体や日常生活の計測をしたこと、など

本実験は、期間が非常に短かったため減量等の身体効果はなかったものの、運動や食事、日常生活、身体に対する意識の高まりや行動の変化が大半の人に現れた。運動や食事習慣はそれを継続することが重要ではあるが、このような機会が意識や行動の変化のきっかけになったことは明らかである。

3. 指導サービスの将来像

本実験では、プロフィール情報の収集とともに、モニターへのセルフモニタリングデータや、簡易のレコメンド情報のフィードバックを行うことで、スポーツジムにおける指導サービスの将来イメージをモニターにある程度意識してもらう工夫を行った。そして実験終了後に事後アンケートの中で今後期待する指導サービスを質問した。本章では、それらの回答を整理し、スポーツジムにおける指導サービスの将来像を描いてみる。

以下に将来サービスのベースとなり得る機能を挙げる。

1. セルフモニタリング

本実験でも既に提供したが、会員から日々収集したデータの変化をグラフなどで分かりやすく表現し会員に返すことで、会員の身体や運動、食事などに対する意識の向上が期待できる。このようなサービスは色々なところで既に始められており、会員の記録の手間を少なくし、さまざまなプロフィール情報が1箇所で見られる環境を整えることが鍵となる。

2. 個人の目的・意識レベル・嗜好等を反映したレコメンド

従来の運動指導などがそうであったように、個人の大きな特性の区分だけで指導内容を決めてきたが、今後はより詳細な個人の特性を反映した指導が必要である。例えば減量や筋力アップ、健康増進、気分転換などスポーツジムに通う目的も個々に異なる。また意識レベルや嗜好、体格などによっても適切な指導内容は異なる。このような個別のレコメンドを行うためには、会員があまり抵抗を感じない方法でプロフィールを収集し、プロフィールからその人の特性を的確に推定することが必要である。

3. 具体性・指標性のあるレコメンド

個人の特性を反映した指導とも重なるが、個人の特性をより詳細に区分するとともに、指導内容がより具体的で、指標性のあることが期待されている。しかし、個々人で厳密には異なる特性ごとに指導内容を用意しておくことは実質不可能であることから、具体性のもたせ方に工夫が必要である。例えば、後述のような、特性が類似する他の人の具体的な行動例を提示するなどの方法が考えられる。

4. 本人が気付いていない生活習慣などの傾向の指摘

普段の生活の中ではなかなか気付くことのない生活習慣などの傾向は、日々のデータを長期的に蓄積して始めて分かることがある。またその人だけのデータを分析しても分からないが他の人との比較をするとその人特有の傾向が明らかになってくることもある。

5. 行動と結果の関係性の提示

現在の行動が今後どのような結果をもたらす可能性があるかを提示することで、意識や行動の改善が促されると

考えられる。行動から結果を推測することは現在は確立した専門的知見をもとに判断されることが多いが、より多くの会員のより詳細な内容の要求に答えるためには蓄積されたデータをもとにルーティン的に行動と結果の関係のルールを発見し提示していく必要がある。

行動と結果の関係性の提示のより具体的なサービスイメージとして異常時のアラート機能が考えられる。過去のデータと他の人のデータをもとに現在の行動にもたらされる結果を推測する。

6. 参考となる他の人の具体的な行動例の提示

具体性のあるレコメンドを作成する方法として、個々のより詳細なレコメンドテンプレートを用意するか、実際の行動事例を利用するか2つの方法が考えられる。後者は協調フィルタリングの1種と考えられるが、Amazon.comの「この本を買った人はこの本も買いました」のような商品が明確な場合は違い、運動指導サービスは推薦商品となる運動内容の境界を定義することが難しい。商品の分類、抽象化が課題となる。

7. コミュニティへの意識機能

他の人の具体的な行動例を提示することとともに、その会員が属しているコミュニティを意識させて上げることで、意欲の向上につながると考えられる。本実験においても、セルフモニタリングデータのフィードバックの際に参加者全員の中での順位を提示してあげる試みを行った。比較的好評であったが、参加者の中にはさまざまな特性の人がいるため同じ基準で順位付けすることは意味がない等の意見もあった。そのため特異性により詳細にグルーピングした上でコミュニティを提示してあげる必要がある。

4. サービス実現のための情報解析の検証

本章では、3.章で挙げたスポーツジムにおける指導サービスの将来条件のうち2、3に対して、今回の実験で収集したようなプロフィール情報をどのように活用すればよいかを検討する。

4.1 個人の特性を反映したレコメンドの作成

本収集実験からも分かったように、個人ごとにその目的(減量、気分転換等)や意識レベル、嗜好、性格などは異なり、それを考慮せずに同じように指導しても効果が現れないことが多い。そのため優れた指導者は会話や行動観察を通してその個人の特性を把握し、その人にあった指導を行っている。しかしこのようにエキスパートにのみ依存した指導サービスではその市場の広がりには限りがあり、今後の多様な需要に答えていくこともできない。よって、従来からある入会時アンケートとともに日々の身体、運動、食事、生活の記録データをもとにその人の特性を分析し、適切なレコメンドを自動的に生成する仕組みを構築していく必要がある。本節では、本実験したプロフィール情報を利用して意識変化に重要な要因が何かを分析し、レコメンド生成につなげる方法を検討する。

2.2節の実験結果でも紹介したように、今回の実験を通じて意識や行動の変化があった人は多く、そのきっかけは1)実際に運動をしたことや2)計測をしたこと、などであることがアンケート結果から分かった。このような要因を日々のプロフィール記録から収集・分析することができると、会員及び提供側の負担なくレコメンドが可能である。

そこで、アンケート結果をもとに意識向上があった人となかった人の区分し、IGVminerを使ってモニタリングデータ

からその要因を分析する。以下に分析方法とその結果をまとめる。

- 利用ツール
IGVcaep (顕在パターンの集約スコア判定法)
- 利用データ
 - 目的変数: 意識向上の有無 (事後アンケート結果より)
 - 説明変数: モニタリングデータから変換された 45 変数 (体重計測欠損数、体重平均、歩数平均、有酸素合計、筋トレ合計、筋トレ頻度、日常運動合計、取り組み度合い平均、性別、目的など)
- パラメータのチューニング
式 (1) に示すように IGVcaep のクラス判定スコアは LCM で抽出された各クラスの顕在パターンの有無によって計算される。しかし LCM が抽出する支持度や増加率が非常に低いパターンも計算に加えては過学習の原因となる。そこで予測精度をある程度に保ちつつ利用パターン数が少ない組み合わせになるような、支持度や増加率の閾値を探索した。その結果、支持度 0.5、増加率 5 を閾値とした場合に、正解率 0.733、未分類率 0.0 となり比較的精度とパターン数のバランスがとれていることが分かった。
- 分析結果
 - 抽出パターン
支持度 0.5、増加率 5 以上の顕在パターンとして、正例 9 パターン、負例 4 パターンが抽出された。正例顕在パターンの例を以下に示す。

(P1) 体重欠損数/少 歩行距離分散/小 主観的強度分散/弱

(P2) 体重欠損数/少 日常運動合計/多 サプリメント合計/少

これらのパターン例からも分かるように、「体重欠損数/少」即ち「計測をきちんと行う」ことや、「日常運動合計/多」即ち「運動量を増やす」ことなどが意識向上の要因となったことが分かる。これはアンケートの結果と一致する。但し、意識変化が先か運動や計測の行動が先かはこれだけでは分からず、おそらく相互に影響し合っていることが事実であろう。

- 個人の特性に合ったレコメンド案の作成

次にこれらの抽出パターンを使って、現状のプロファイル特性に対してどの項目を変えれば意識向上につながるかを分析した。方法は説明変数のうちの 1 つの値を変更することで、正例スコアが上がり負例スコアよりも高くなるかを判断した。具体的な向上事例を示す。

(事例 1) A さんの場合

現状: 体重欠損数/多 (正スコア:0、負スコア:1.0)

改善案: 体重欠損数/少 (正スコア:0.6、負スコア:0.5)

(事例 2) B さんの場合

現状: 日常運動合計/少 (正スコア:0.6、負スコア:0.4)

改善案: 日常運動合計/多 (正スコア:1.2、負スコア:0.4)

これらの例から分かるように、日々の計測が少ない人に対して計測を促したり、運動量の少ない人により多くの運動をするように促すことで意識向上につながることを期待できる。

4.2 具体性のあるレコメンドの作成

前節では意識向上のために「日々計測を行う」、「運動量を増やす」のような非常にメタなレベルのレコメンド案を自動生成した。しかし、実際の指導の場面では単に運動量を増やさないと言っても具体的にどのような運動をどのような時間・場所で増やせばよいか分からない。そのため、従来は大まかな特性区分に対してレコメンドのテンプレートを用意することで対応してきた。しかし、さまざまな特性ごとにきめの細かい指導を行うためにはテンプレートを用意するだけでは限界がある。そのため、我々は本人の過去の行動や他の人の行動を参考事例として推薦することを考えている。本稿では紙面の関係からその考え方だけを示すことにとどめる。

3. 章の 6 で既に述べたように、運動や食事内容を推薦する上でその内容の境界を定義することが課題となる。例えば、「A さんは 1 月 26 日 13 時に B 店にてランニングマシンを使って有酸素運動を 1 時間やった」という事例があった場合にこれをそのまま推薦することは不必要な具体性であり、また個人のプライバシーからも問題がある。そのため、レコメンドの問題には具体化とともにその逆の抽象化の課題がある。このバランスを制御して大量に機械的なレコメンド案を作成するためには、これまであまり明確な基準もなくプロフィール情報の具体化・抽象化を行っていたものに基準を設ける必要がある。例えば、「A さんは～」のような場合、1) 人、2) 時間、3) 場所、4) 行為内容、5) 分量、の 5 つの抽象化が考えられる。プロフィール情報を収集し、共通パターンの抽出、レコメンドの具体化の一連の流れにおいては、この抽象化の基準をもとに具体例を一旦抽象化し、共通パターンを導出し、次に抽象化の逆方向に下ることでレコメンドを具体化することで、機械的にレコメンドを生成することができる。

5. まとめ

本稿では、情報大航海プロジェクトの中で我々が関わったプロフィール情報解析基盤 IGVminer を紹介した。そしてその利用実証実験としてスポーツジムにおけるプロフィール収集実験と、収集されたデータをもとにユーザの特性を推定し適切な運動指導を行うための方法を提案した。今後はより本格的にサンプル数を増やし身体効果が期待できる期間長でプロフィール収集を行っていききたいと思う。また、データマイニング技術などを用いていかに指導サービスの将来像で挙げた機能を実現していくかを検討していきたい。

謝辞

本開発研究は情報大航海プロジェクトの中で実施された。関係者の皆様に感謝する。また実験の場を提供して下さった (株) 東急スポーツオアシス殿、実験に参加頂いたモニターの皆様にも感謝する。

参考文献

- [Uno03] T.Uno, Y.Uchida, and H.Arimura: LCM: An efficient algorithm for enumerating frequent closed item sets, Proc. of IEEE ICDM'03 Workshop FIMI'03, 2003.
- [湊 05] 湊真一: VSOP: ゼロサプレス型 BDD に基づく「重み付き積和集合」計算プログラム, 電子情報通信学会技術研究報告, 105(72) pp.31-38, 2005.
- [Dong99] G.Dong, X.Zhang, L.Wong, and J.Li : CAEP: Classification by Aggregating Emerging Patterns, Proc. of 2nd International Conf. on Discovery Science, 1999.