

ウェアラブル環境における行動解析のためのイベント生成手法

6E-7

大田 智数<sup>†</sup> 波多野 賢治<sup>†</sup> 吉川 正俊<sup>†,‡</sup> 植村 俊亮<sup>†</sup>

<sup>†</sup> 奈良先端科学技術大学院大学 情報科学研究科 <sup>‡</sup> 国立情報学研究所 ソフトウェア研究系

1 はじめに

近年、衣類のように日常的に身につけて利用するウェアラブルコンピュータの研究が盛んである [1]。また、地理情報の電子化や GPS(Global Positioning System) などの位置計測技術の発展により、屋外で地理情報システム (GIS:GeographicInformation System) を利用できる環境も整いつつある。こうした背景から、利用者の現在地に関連した情報を提供するためのサービスとして位置情報サービスが注目されている。

本稿では、ウェアラブルコンピュータに装備された GPS などの位置計測機器を利用して現在地を容易に特定することができ、位置情報サービスを利用できる環境の事をウェアラブル環境と言う。ウェアラブル環境では利用者の日常生活における移動経路を GPS データ (緯度, 経度) や現在地データの履歴として残すことが可能となる。現在地データとは GPS データの中でも特に利用者が街中の商店や施設に近付いた時の GPS データを、その店名や施設名といった現在地の名称で表現したデータである。そのため現在地データの履歴は、GPS データの履歴による詳細な移動経路を簡略化したものとなる。我々は以前に、GPS データの履歴からイベントを生成し、時間依存の相関ルールであるエピソードルール [2] を適用して、行動パターンを発見するための手法を提案した [3]。本稿では、現在地データを GPS データの中でも特に意味のあるデータとして考え、GPS データと現在地データの履歴から生成されるイベントを比較検討した結果について報告する。

2 行動パターンの解析

2.1 システムの概略

本システムの処理の流れを図 1 に示す。

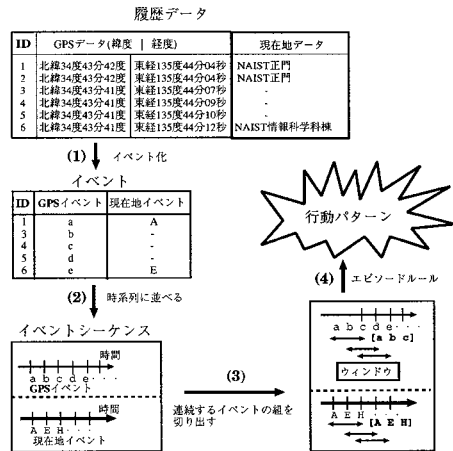


図 1: システムの処理手順

図 1 の(1)では、まずウェアラブル環境で取得できる GPS データや現在地データで構成される利用者の移動した履歴からイベントを生成する (2.2 節に後述)。次にイベントを時系列に並べてイベントシーケンスを生成する (図 1 の(2))。さらに、ウィンドウと呼ばれる近傍距離を導入し、イベントシーケンスから一定範囲 (ウィンドウサイズ) でイベントの組 (エピソード) を切り出し、ウィンドウを時間軸に対して正の方向に 1 ずつシフトすることで、全てのエピソードを抽出する (図 1 の(3))。本稿ではイベント a, b, c から構成されるエピソードを [abc] と定義する。[abc] は a の後に b, その後に c が起きるといった順序関係を反映したイベントの組となる。最後に図 1 の(4)において、各エピソードの頻度 (freq), 支持度 (support), 確信度 (confidence) という評価値を求め、閾値よりも高いエピソード (エピソードルール) を行動パターンとする。エピソード [abc] の支持度, 確信度はそれぞれ以下の式で求める。

$$support([abc]) = \frac{freq(a \cap b \cap c)}{freq(all\ episodes)}$$

A Method of Event Generation for Behavior Analysis Using Wearable Computer  
Tomokazu Ohta<sup>†</sup>, Kenji Hatano<sup>†</sup>, Masatoshi Yoshikawa<sup>†,‡</sup>, Shunsuke Uemura<sup>†</sup>  
<sup>†</sup> Graduate School of Information Science, Nara Institute of Science and Technology  
<sup>‡</sup> Software Research Division, National Institute of Informatics

$$\text{confidence}(\{abc\}) = \frac{\text{freq}(a \cap b \cap c)}{\text{freq}(a \cap b)}$$

支持度はイベントシーケンス上のエピソードの割合を意味し、確信度は a, b の後に c が続く割合を表している。

## 2.2 イベントの生成

本稿では、利用者の通過した場所を以下の方法で記号化したものを現在地イベントと言う。まず、現在地データの履歴を時系列に並べる。このとき、名称が異なる現在地データを異なる記号で表現し、名称が同じ現在地データは同一の記号で表現する。また、同じ記号が連続する場合は、ひとつの記号にまとめてしまう。そうすることでデータの冗長性を無くし、重複する記号で構成されるエピソード [aa] における評価式の矛盾を無くすことができる。このように現在地データの履歴を記号化したものを現在地イベントと呼ぶ。GPS イベントは、現在地イベントが名称の違いで記号化したのに対して、緯度、経度の違いから記号化することで、同様に記号化することができる。本稿では GPS データから生成したイベントを GPS イベントと呼ぶ。

## 3 実験と考察

GPS イベントと現在地イベントを比較するために、第一著者の日常生活から取得した通学などの GPS データ 1 ヶ月分に対してデータマイニングを行った。実験には Sony のハンディ GPS レシーバ PCQ-GPS33SUG を使用しサンプリングレート 1 秒、誤差 30m 程度で GPS データを取得し、GPS イベントを生成した。現在地データは、国土交通省が配布している国土数値情報の公共施設を利用し、店名などは手作業で GPS データと対応させてデータを取得し、この履歴から現在地イベントを生成した。GPS イベントと現在地イベントからルールを導出した結果を図 2 に示す。

図 2 は頻度 1 以下のものをノイズとして除去したエピソードに対して確信度が 0.5 以上のものをエピソードルールとして導出している。この図からウィンドウサイズ 3 のとき、両方のイベントで全エピソードに対するルールの割合が最大になっているのが分かる。ここでルールの割合が高いということは、無駄なイベントが少ないことを意味する。ウィンドウサイズが小さい程、ルール数が増えることが予測されるが、このとき確信度は反対に低くなる傾向がある。そのため、ウィンドウサイズ 2 のときよりも、ウィンドウサイズ 3 の方がルール数が高くなっていると考えられる。こ

の結果からウィンドウサイズ 3 が各ウィンドウサイズの中で最適であると考えられる。また、ウィンドウサイズ 3 のとき、現在地イベントの方が GPS イベントに比べ、ルールの割合が高い値を示している。現在地データは GPS データの部分集合であるが、この結果から元データである GPS データにおけるノイズの割合が現在地イベントでは少なくなっている事が分かる。以上の結果から現在地イベントの方が GPS イベントに比べ、ノイズもデータ量も少ない効率的なイベントと言える。そして、現在地イベントが GPS イベントの中でも特に意味のあるデータを取り出す事に成功していると考えられる。

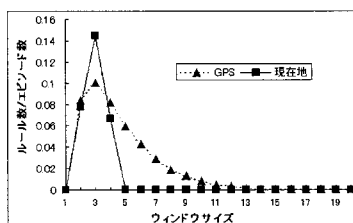


図 2: GPS イベント、現在地イベントから生成されるルールの割合

## 4 おわりに

本稿ではウェアラブル環境において取得できる GPS と現在地の履歴を用いて、移動する利用者の行動を解析するためのイベント生成手法について述べた。今後は、滞在時間や徒歩、車といった異なる移動手段をエピソードの抽出に反映したウィンドウを実装する予定である。

謝辞 本研究の一部は、文部科学省科研費基盤研究 (11480088, 12680417, 12208032) および科学技術振興事業団戦略的基礎研究推進事業の支援を受けている。ここに記して謝意を表します。

## 参考文献

- [1] Bradley J. Rhodes, Nelson Minar, and Josh Weaver. Wearable computing meets ubiquitous computing reaping the best of both worlds. In *Proceeding of the 3rd International Symposium on Wearable Computers*, pp. 141-149, 1999.
- [2] Heikki Mannila, Hannu Toivonen, and A.Inkeri Verkamo. Discovery of frequent episodes in event sequences. *Data Mining and Knowledge Discovery*, Vol. 1, pp. 259 - 289, 1997.
- [3] 大田智数, 波多野賢治, 吉川正俊, 植村俊亮. ウェアラブル環境における行動履歴を用いた情報提示. 電子情報通信学会技術研究報告, Vol. 101, No. 110, DE2001-1, pp. 1-7, June 2001.