クラスタリングによる人間の振舞い認知

正員 小越 康宏 (富山大学)

非会員 小越 咲子 (金沢大学)

非会員 広瀬 貞樹 (富山大学)

非会員 木村 春彦 (金沢大学)

Recognition of Human Activities Using Clustering Analysis

Yasuhiro OGOSHI, Member (Toyama University), Sakiko OGOSHI, Non-member (Kanazawa University), Sadaki HIROSE, Non-member (Toyama University), Haruhiko KIMURA, Non-member (Kanazawa University)

Recently, the number of the aged person who leads a solitary life is increasing, and unexpected accidents in a residence have been increasing rapidly. Therefore, understanding of their activities in their residences will be very important to avoid accidents and to take care of them. In this paper, the new method to set up a sensor of which signals are transferred to the computer in the room and to analyze their activities using the new method is described.

キーワード：赤外線センサ，クラスタリング，遺伝的アルゴリズム，データマイニング，振舞い認知
Keywords: infrared sensor, clustering, genetic algorithms, datamining, human activities

1. はじめに

独居老人が家庭内で不慮の事故に遭遇したり，体調に異変が生じた場合，発見が遅れ死に至る深刻なケースがあるので，異変を早期に発見することは非常に重要である。

著者らは，赤外線センサで人体を感知し，振舞い認知を試みてきた⁽¹⁾。時間帯別に滞在時間の傾向を調べ，これを通常の振舞いとして記録し，観測データがこれよりも長くなる場合は異常と判別する手法⁽²⁾を提案したが，振舞いのパターン化はなされていない。そこで，本論文では，トイレにおいて，入室時刻と滞在時間の2変量についてクラスタリングし，正規分布モデルに当てはめることにより振舞いとして捉える手法を提案する。

2. 振舞い認知の方法

以下に，提案した振舞い認知の手法について説明する。

〈2・1〉振舞いデータの計測 赤外線センサをトイレに設置し，人体の検出信号（1：存在する，0：存在しない）を基に，[入室時刻，退室時刻，滞在時間]の形式でデータを時系列保存する。

〈2・2〉振舞いデータの分類 保存されたデータから，入室時刻と滞在時間の2変量に対しクラスタリングする。クラスタリングには，計算回数を少なくするために，以下に示す遺伝的アルゴリズム⁽³⁾（以下GAと記す）を用いた。

〈2・2・1〉GAを用いたクラスタリング処理 データの番号を k ($k = 1, 2, \dots, n$)とし，データを x_k とする。 x_k は， x_k^1 ：入室時刻， x_k^2 ：滞在時間の二つの要素からなる。

データを c 個にクラスタリングすることを考える。クラスタの番号を i ($i = 1, 2, \dots, c$)とし，クラスタ i に属するデータ x_k の重心を g_i とする。この g_i も，同様に g_i^p ($p = 1, 2$)の二つの要素からなる。

データ x_k がクラスタ i に属するかどうかを，2値変数 u_{ik} （1：属する，0：属さない）で表す。

クラスタ i に属するデータ数を n_i とし，この値と重心 g_i を式(1)，(2)に従って求める。

$$n_i = \sum_{k=1}^n u_{ik} \quad \dots\dots\dots (1)$$

$$g_i^p = \frac{\sum_{k=1}^n u_{ik} x_k^p}{n_i} \quad (p = 1, 2) \quad \dots\dots\dots (2)$$

データ x_k の属するクラスタの番号を i_k とすると，遺伝子型を式(3)のように与える。

$$(i_1, i_2, i_3, \dots, i_n) \quad \dots\dots\dots (3)$$

また，GAの各世代における適応度を E とおき，以下の評価関数(4)を用いる。

$$E = \frac{\sum_{l=1}^{c-1} \sum_{m=l+1}^c \|g_m - g_l\|}{\sum_{k=1}^n \sum_{i=1}^c u_{ik} \|x_k - g_i\|^2} \quad \dots\dots\dots (4)$$

この式は，効率的にクラスタリングを行うために，各クラスタの重心間の距離が離れるほど分子は大きく，各クラスタ内のデータ間の距離が近いほど分母の値が小さくなるようにした。

また，データ x_k の各要素 x_k^1, x_k^2 については，以下のよう

表1 トイレにおける5日間の観測データ

Table 1. Human activities for 5 days (Rest room).

Data No.	Entrance Time [Hour:Min:Sec]	Exit Time [Hour:Min:Sec]	Duration [Min]	(Activities)
1	06:03:48	1,06:05:34	1.77	Urine
2	07:49:01	1,07:53:24	4.38	Stool
3	10:27:34	1,10:28:38	1.07	Urine
4	:	:	:	:
7	21:22:14	1,21:23:29	1.25	Stool
8	:	:	:	:

に正規化してからクラスタリング処理を行う。

- x_k^1 : 時刻0時~24時を0~1に対応
 - x_k^2 : 各滞在時間を最大滞在時間で除算
- GAのパラメータは以下のように設定した。
- 個体総数: 50
 - 交差の生起確率: 0.1
 - 突然変異率: 0.01

このようにパラメータを設定した理由を以下に示す。

個体総数を大きくすれば多様性を維持できるが、計算のステップ数が多くなるので、計算のステップ数をなるべく少なくするように、個体総数を50とした。

交差の生起確率を大きくすればより良い組み合わせが見える可能性があるが、あまり大きくすると収束が妨げられるので、交差の生起確率を小さくとり0.1とした。

突然変異はそれ自体が強力なオペレーションであり、作用を大きくするとランダムサーチに近くなり収束が妨げられるので、突然変異率を小さくとり0.01とした。

〈2・3〉 振舞いデータのモデル化 入室時刻、滞在時間の2変量について〈2・2〉で処理された各クラスタに対して、以下の手順で正規分布モデルに当てはめる。

〈2・3・1〉 クラスタごとの正規分布モデル化 データ x_k の各要素 x_k^1, x_k^2 の標本平均を $\bar{x}_i^p = g_i^p$, 標本標準偏差を s_i^p とおき (5) に従い計算する。

$$s_i^p = \sqrt{\frac{\sum_{k=1}^n (u_{ik}(x_k^p - \bar{x}_i^p))^2}{n_i - 1}} \quad (p = 1, 2) \cdots (5)$$

これらの値に従う正規分布 $N(\bar{x}_i^p, (s_i^p)^2)$ を求める。

〈2・3・2〉 クラスタごとの信頼区間設定 〈2・3・1〉においてモデル化された各クラスタの正規分布に対し、それぞれ95%の信頼区間を求め、これらの区間を各クラスタの一般的な範囲と考える。

3. 実験

2の手順に従って実験を行い、正規分布モデルに当てはめた結果を以下に示す。センサからの検出結果は表1のようなデータベース形式で保存される。この例では、参考までに実際の行動 (Activities) も示してある。

クラスタリング[†]を行った結果を図1に示す。図1の横軸

[†] トイレにおける振舞いは、滞在時間で容易に大便と小便のクラスに分けることができるが、トイレ以外の部屋で複雑な振舞いのクラス分けをするにはクラスタリングが有効と考えられる。例えば、ガスセンサ⁽⁴⁾等の情報を融合して、クラス分けを同定することも検討している。

図1 トイレにおけるクラスタリングされたデータと正規分布における95%の信頼区間

Fig. 1. Clustering data on the rest room and 95% of reliability range in normal distribution.

は入室時刻を、縦軸は滞在時間を表す。実験において、最大滞在時間は4.38分となり、この値を用いて縦軸を正規化した。図1におけるクラスと、実際の振舞い(表1の観測データ(行動))と照らし合わせて見ると、下側の○で示された点が小便の行動を表し、上側の△で示された点が大便の行動を表しており、このデータにおいては各振舞いを分類することができた。このように、観測されたデータをクラスタリングすることにより振舞いを捉えることができると考えられる。

正規分布でモデル化し、95%の信頼区間を求めると図1のように、長方形で囲んだような範囲となる。

4. むすび

今回は1人のトイレの振舞いについて、本手法の有効性を確認したが、振舞いは個人差があり、有効性を検証するために、より多くの人物に対しても実験を行い、各人の振舞いを比較する必要がある。また、振舞いを長方形のクラスタとして捉えているが、他のクラスタリング手法を用いて、より柔軟なクラス分けを行ったり、ふさわしいクラスタの形状も検討する必要があると考えられる。

謝辞 査読委員の先生から貴重な御教示を頂いた。ここに記して謝意を表する。本研究は財団法人ユニバーサル財団研究助成による成果の一部である。

(平成14年4月17日受付, 同14年5月17日再受付)

文 献

- (1) Y. Ogoshi, T. Misawa, H. Kimura, Y. Ueda, and T. Oyabu: "Recognition of human behavior using sensor-agent system," Proc. The Third Asia-Pacific Conference on Industrial Engineering and Management Systems, pp.214-218, Hong Kong (2000)
- (2) 小越康宏, 小越咲子, 広瀬貞樹, 木村春彦: "赤外線センサ情報からのデータマイニングによる独居老人の振舞い認知に関する一考察," 信学論 (D-II), Vol.J85-D-II, No.5, pp.959-964, (2002)
- (3) D. E. Goldberg: "Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, Massachusetts (1989)
- (4) 大藪 多可志, 木村 春彦: "ガスセンサによる人間のふるまい認知," 電気学会計測研究会, IM-94-11, pp.77-85, (1994).