

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 16 日現在

機関番号：32305

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500813

研究課題名(和文)健康データマイニングによる個人の特質クラスタ解析に関する研究

研究課題名(英文)Cluster Analysis of Characteristics in the Nature of Personal Health with Healthcare-Data-Mining

研究代表者

竹内 裕之(TAKEUCHI, HIROSHI)

高崎健康福祉大学・健康福祉学部・教授

研究者番号：00348129

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：個人の日々のエネルギー消費・摂取量と体脂肪率変化の時系列相関解析において、最大の相関係数をもたらすエネルギー消費・摂取量の時系列加算処理の重み付パターンを求めた。重み付パターンは2つのパラメータで表現され、その組み合わせが個人の健康に関する特質を表すとの想定のもとに解析を行った結果、(1)ある個人の長期的な時系列データにおいて、加齢や季節に関係なくパラメータの組み合わせはある偏りを示す、(2)複数の被験者の時系列データでは、パラメータの組み合わせは大きく2つのグループに分かれる傾向がある、ことなどが判った。

研究成果の概要(英文)：Correlations between energy expenditure/supply and body-fat percentage were studied using personally stored daily time-series data. The weighting patterns for the summation of daily time-series energy expenditure and supply data giving the maximum correlations with the variation of daily body-fat percentage data were obtained. The weighting patterns can be expressed by two parameters whose combination is considered to characterize the nature of personal health. The combination of the parameters for a subject was found to show a significant bias in the frequency distribution, independent of season and aging, for the term of seven years, and the combination of the parameters of 20 other subjects showed a tendency to divide into two types.

研究分野：健康情報学

科研費の分科・細目：健康・スポーツ科学 応用健康科学

キーワード：健康データマイニング 個人健康管理 時系列データ解析 遅延相関分析法 クラウドコンピューティング

1. 研究開始当初の背景

(1) 健康医療分野の情報化は、医療機関を中心とした従来の情報化の波に加え、e-healthというキーワードに象徴されるように健康分野で大きな進展をみせている。特にインターネットを存分に活用するクラウドコンピューティングの応用分野として、健康情報サービスが新たな局面を迎えている。

(2) 個人の生活習慣や、体重、体脂肪率など日常家庭でも取得できる健康データをクラウド上に蓄積し、個人の健康管理に役立てようとする試みが各所でなされているが、蓄積されたデータの二次利用という面では研究開発が十分に行われていない。

2. 研究の目的

(1) 携帯電話を情報端末として、個人の生活習慣と日常の健康に関するデータをインターネット上のサーバに蓄積し、サーバ上で生活習慣と健康状態の相関ルール抽出(健康データマイニング)を行って個人に通知する、クラウドコンピューティングをいち早く具現化した「個人健康管理システム」および「特定健診・保健指導を支援するシステム」を開発してきた。本研究では、開発したシステムを活用し、生活習慣と健康状態の相関における個人差を特徴付けるパラメータを抽出して、健康に関する個人の特質のクラスタ解析を試みる。

(2) 上記個人差を踏まえ、蓄積されたデータの二次利用の有力な手法として開発した健康データマイニングにより自動抽出された生活習慣と健康状態に関わるルールやパターンの妥当性を検証する。

3. 研究の方法

(1) 我々が開発している健康データマイニングでは、「生活習慣の蓄積が健康状態に変化をもたらし、その影響は時間遅れをもって現れることがある」という極めてシンプルなモデルをベースとしている。すなわち、ある健康状態(例えば体脂肪率)の変化を目的変数とし、時間遅れを考量したある期間の生活習慣(例えば運動による消費エネルギー)の蓄積を入力変数として相関ルールを抽出する。この手法を遅延相関分析法と呼ぶ。

(2) 本研究では、遅延相関分析法をベースとして、データの時系列相関解析に、新たに重み付の概念を導入した。これまでは、ある期間の生活習慣データを単純加算し、加算日数と遅延期間をパラメータとして健康状態の変化との相関を解析していた。しかし、生活習慣データを過去に遡って毎日に単純加算するというのは、健康状態との関わりにおいて複雑な生体反応をあまりにも単純化していると考えられる。そこで、過去の生活習慣データを毎日に適当な重みを付けて加算

し、その最適な重み付パターンそのものが個人を特徴付ける指標になると考えた。具体的には重み付に正規分布関数を活用し、分布を特徴付ける2つのパラメータ μ と σ の組み合わせで個人差を特徴付けることを試みた。

4. 研究成果

(1) 生活習慣データの時系列加算処理に μ と σ の2つのパラメータを持つ正規分布関数型を活用した重み付けの概念を導入し、健康状態の変化と生活習慣データの時系列加算の間に最大の相関をもたらし重み付けのパラメータの組み合わせ(μ, σ)が個人の特質を反映しているという想定のもとにデータ解析を行った結果を以下に纏めた。

1 被験者の7年余りに亘る体脂肪率、消費カロリーおよび摂取カロリーの時系列データを3カ月毎に区切って解析した結果、体脂肪率と消費カロリー、摂取カロリーそれぞれの関係において最大の相関係数を示す(μ, σ)の組み合わせが特定の組み合わせ範囲に偏っている(図1、図2)。

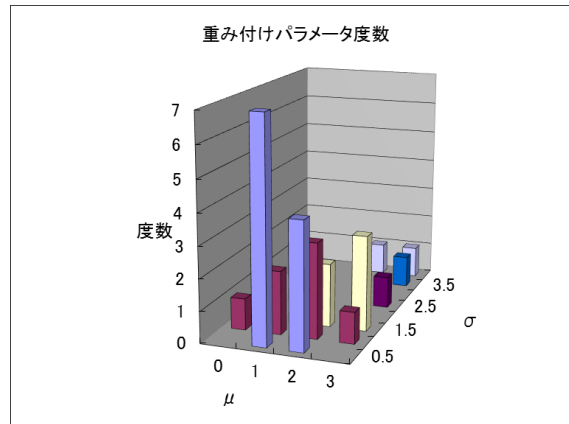


図1. 消費カロリーと体脂肪率の時系列データ間の相関係数が最大となる(μ, σ)のセットの度数分布(被験者1名、7年間計27セット)

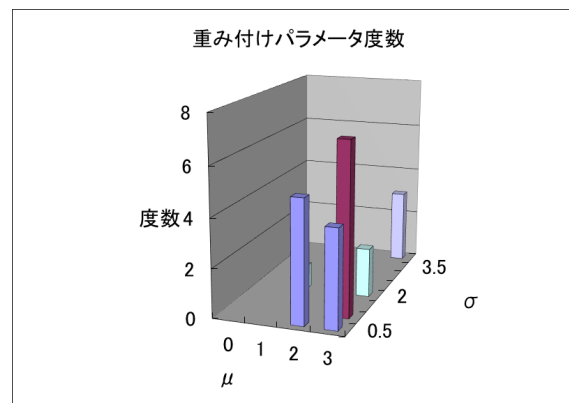


図2. 摂取カロリーと体脂肪率の時系列データ間の相関係数が最大となる(μ, σ)のセットの度数分布(被験者1名、7年間計22セット)

その偏りは、加齢や季節に関係がないことから個人の特質を反映している可能性がある。

20 被験者の半年間の体脂肪率と消費カロリーの時系列データを3カ月毎に区切って解析した結果、(a)低 (μ も小さい)グループと、(b)高 μ (σ は広く分布)グループに分かれる傾向がみられる(図3)。

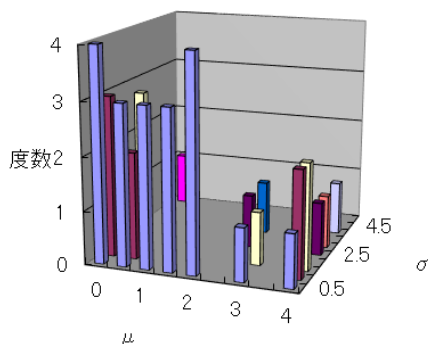


図3.消費カロリーと体脂肪率の時系列データ間の相関係数が最大となる(μ, σ)セットの度数分布(被験者20名、各2セット計40セット)

(2)生活習慣と健康状態の相関に関わる個人差のクラスタ解析を行う場合に、個人の長期に亘る健康データの変動そのものにも個人の特質が現れることに着目する必要があることが明らかになった。以下に8年余りに亘る1ユーザのデータ解析結果を纏めた。

最大血圧(収縮期血圧)、最小血圧(拡張期血圧)とも顕著な生理学的季節変動を示し、夏冬の変動幅は、疫学的研究で得られている値(最大血圧で10 mmHg、最小血圧で7 mmHg)より大きく、それぞれ約20 mmHg、約10 mmHgであった。

前記変動幅は、特に最大血圧において年々大きくなる傾向があり、これは先行研究と照らし合わせ加齢による変化と考えられる。

血圧ほど顕著ではないが、脈拍数においても生理学的な季節変動がみられ、冬場は夏場に比べ平均約10回/分多かった。

体脂肪率も顕著な季節変動を示した。これは自然環境下に生きる動物と同じように気温の低下とともに生理反応として貯めこんだ脂肪を現代のヒトは冬に消費することなくキープしてしまう結果と考えられている。本研究の対象者にみられる変動幅の年々の増加は、生理反応として貯めこんだ脂肪を冬に消費する機会が加齢とともに益々減ることの結果と考えられる。

(3)本学の学生を中心とした個人健康管理システムのボランティアユーザが、2012年6月1日から11月30日までの6か月間に日毎の粒度で蓄積した生活習慣と健康に関するデータに基づき、開発した健康データマイニング手法によって得られたパターンやルールについてその妥当性を検証した。その結果を以下に纏めた。

本研究で対象とした3ボランティアユ

ーザの内2名についてはサーバに蓄積された時系列データに基づき、開発した健康データマイニングのアルゴリズムによって、適切に生活習慣と健康状態の間の相関ルールが自動抽出されていた。

他の1名については、誤りではないが妥当ではない相関ルールが自動抽出され、その原因は生活習慣データと健康状態の間に強い非線形の相関があることによることが判った。

上記の場合、生活習慣データの範囲に制限を設けることにより適切に健康状態との間の相関ルールが抽出されることを確認できた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計10件)

竹内裕之、児玉直樹、健康データマイニングの評価()-6か月間の時系列データに基づく事例-、高崎健康福祉大学紀要、査読有、第13号、2014、pp.1-8

竹内裕之、児玉直樹、健康データマイニングにより自動抽出されたルールの妥当性検証、第6回データ工学と情報マネジメントに関するフォーラム論文集、査読無、2014、F6-4 pp.1-6、

<http://db-event.jpn.org/deim2014/final/proceedings/F6-4.pdf>

Hiroshi Takeuchi, Yuuki Mayuzumi, Naoki Kodama, and Keiichi Sato, Personal Healthcare System Using Cloud Computing, Proc.14th World Congress on Medical and Health Informatics (Medinfo 2013), 査読有、2013、pp.936, DOI:10.3233/978-1-61499-289-9-936

竹内裕之、黛 勇気、児玉直樹、健康と生活習慣に関わる時系列データ解析に基づく p-health の1例、高崎健康福祉大学紀要、査読有、第12号、2013、pp.11-19

竹内裕之、黛 勇気、児玉直樹、個人の日常の健康と生活習慣に関わる時系列データ取得の意義、第5回データ工学と情報マネジメントに関するフォーラム論文集、査読無、2013、E5-6 pp.1-4、

<http://db-event.jpn.org/deim2013/proceedings/pdf/e5-6.pdf>

Hiroshi Takeuchi, Yuuki Mayuzumi, and N. Kodama, Parameters Characterizing Nature of Personal Health in the Correlation between Energy Expenditure/Supply and Body-Fat, Proc. 34th Annual International Conference of the IEEE EMBS, 査読有、2012、pp.2140-2143, DOI:10.1109/EMBC.2012.6346384

竹内裕之、黛 勇気、児玉直樹、個人の生活習慣と健康状態の相関を特徴付ける

パラメータの解析、高崎健康福祉大学紀要、査読有、第 11 号、2012, pp.15-22
黛 勇気、竹内裕之、児玉直樹、佐藤恵一、クラウドコンピューティングによる個人健康管理、第 4 回データ工学と情報マネジメントに関するフォーラム論文集、査読無、2012, D11-2 pp.1-6,
<http://db-event.jpn.org/deim2012/proceedings/final-pdf/d11-2.pdf>
竹内裕之、黛 勇気、児玉直樹、個人のエネルギー消費・摂取量と体脂肪率の相関を特徴づけるパラメータに関する研究、第 4 回データ工学と情報マネジメントに関するフォーラム論文集、査読無、2012, D11-3 pp.1-5,
<http://db-event.jpn.org/deim2012/proceedings/final-pdf/d11-3.pdf>
Hiroshi Takeuchi, Yuuki Mayuzumi, and Naoki Kodama, Analysis of Time-Series Correlation Between Weighted Lifestyle Data and Health Data, Proc. 33rd Annual International Conference of the IEEE EMBS, 査読有, 2011, pp.1511-1514,
DOI:10.1109/IEMBS.2011.6090345

〔学会発表〕(計 6 件)

竹内裕之、児玉直樹、健康データマイニングにより自動抽出されたルールの妥当性検証、DEIM 2014、2014 年 3 月 4 日、兵庫県淡路島夢舞台国際会議場

竹内裕之、インターネットを用いた個人健康情報の蓄積と健康管理、第 8 回日本禁煙科学会学術総会 in ぐんま、2013 年 11 月 3 日、群馬県高崎市高崎健康福祉大学

竹内裕之、黛 勇気、児玉直樹、個人の日常の健康と生活習慣に関わる時系列データ取得の意義、DEIM 2013、2013 年 3 月 4 日、福島県郡山市磐梯熱海ホテル華の湯

Hiroshi Takeuchi, Yuuki Mayuzumi, and N. Kodama, Parameters Characterizing Nature of Personal Health in the Correlation between Energy Expenditure/Supply and Body-Fat, 34th Annual International Conference of the IEEE EMBS, 29 August, 2012, San Diego, California USA

黛 勇気、竹内裕之、児玉直樹、佐藤恵一、クラウドコンピューティングによる個人健康管理、DEIM 2012、2012 年 3 月 5 日、兵庫県神戸市シーサイドホテル舞子ピラ神戸

竹内裕之、黛 勇気、児玉直樹、個人のエネルギー消費・摂取量と体脂肪率の相関を特徴づけるパラメータに関する研究、DEIM 2012、2012 年 3 月 5 日、兵庫県神戸市シーサイドホテル舞子ピラ神戸

〔図書〕(計 1 件)

達 吉郎、竹内裕之 他、(株)エヌ・テ

ィー・エス刊、パーソナル・ヘルスケアーユビキタス・ウェアラブル医療実現に向けたエレクトロニクス研究最前線、2013 年、pp.147-154

6. 研究組織

(1) 研究代表者

竹内 裕之 (TAKEUCHI HIROSHI)
高崎健康福祉大学・健康福祉学部・教授
研究者番号：00348129

(2) 研究分担者

倉林 正彦 (KURABAYASHI MASAHIKO)
群馬大学・医学系研究科・教授
研究者番号：00215047
児玉 直樹 (KODAMA NAOKI)
高崎健康福祉大学・健康福祉学部・准教授
研究者番号：50383146