

クラウドコンピューティングによる個人健康管理

黛 勇気 竹内 裕之[†] 児玉 直樹[†] 佐藤 恵一[‡]

高崎健康福祉大学大学院 健康福祉学研究科 〒370-0033 群馬県高崎市中大類町 37-1

[†]高崎健康福祉大学 健康福祉学部 医療情報学科 〒370-0033 群馬県高崎市中大類町 37-1

[‡]株式会社日立ソリューションズ 〒230-0045 神奈川県横浜市鶴見区末広町 1-1-43

E-mail: {0910404, [†]htakeuchi, [†]kodama}@takasaki-u.ac.jp, [‡]keiichi.sato.jw@hitachi-solutions.com

あらまし 近年ではクラウドコンピューティングを活用した健康管理が大きな潮流になっている。我々はこれまでに、携帯電話や PC を活用した個人健康管理システムや特定保健指導を支援するシステム（健康管理 ASP システム）を開発してきた。これらは、ユーザごとに蓄積された生活習慣および健康状態に関する時系列データに基づき、両者の相関ルールを抽出（健康データマイニング）し、健康管理や保健指導に役立てることを目的としたシステムである。この健康管理 ASP システムを用いて、16 名の被験者を対象に相関ルールの抽出を試み、6~7 割のユーザで洞察を得ることができた。

キーワード クラウドコンピューティング, 健康データマイニング, 時系列データ解析, 生活習慣, 健康状態, 相関ルール

1. はじめに

1.1. 健康管理と情報技術

わが国の医療費は増大をつづけている[1]。中長期的な視野に立つと、メタボリックシンドロームに代表される生活習慣病を予防することが、国民医療費を抑制することにつながる。そのため、生活習慣病に関する国民の意識を高めることが肝要である。

近年では、ICT（Information and Communication Technology）技術を活用したクラウドコンピューティングによる健康管理が大きな潮流になっており、健康医療分野では、民生用の血圧計や体重計、歩数計などを USB（Universal Serial Bus）や BT（Bluetooth）で PC（Personal Computer）や Internet などに接続することで、得られたデータを PC やネットワーク上の DB（Data Base）サーバに蓄積することが可能な機器もある。健康医療分野の情報化は、従来の医療機関を中心とした情報化の波に加え、昨今の医療情報に関する国際学会の潮流として、m(mobile)-health や p(personalized)-health といった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、これらの概念が浸透してきている[2]。特にインターネットを活用す

るクラウドコンピューティングの一つの応用分野として、健康情報サービスは新たな局面を迎えている。

1.2. 個人健康管理システム

我々は健康管理における ICT の利活用いち早く着目し、これまでに健康データマイニングとして個人について「生活習慣の積み重ね」と「健康状態の変化」との関係についての洞察を得るための研究を進めてきた。健康データマイニングは、個人ごとに生活習慣と健康状態に関する時系列データを解析して両者の間の相関ルールを抽出する技術である。これを自動的にインターネット上でサービスする、個人健康管理システム（Personal Dynamic Healthcare System：PDHS）を開発・試行してきた[3,4]。このシステムでは、携帯電話（Mobile Phone）を情報端末として、個人が自分自身の生活習慣と健康状態のデータをインターネット上の DB サーバに蓄積する。一定数のデータが蓄積されると、自動健康データマイニングのサービスがユーザごとにデータを解析し、その個人にユニークな相関ルールを抽出し、結果をユーザに通知するというシステムである。

2. 研究方法

2.1. 健康管理 ASP システム

本研究に用いた健康管理 ASP システムについて述べる[5].

個人健康管理システムをカスタマイズしたシステムを用い、個人ごとに抽出された相関ルールを根拠にした保健指導について、共同研究先（調剤薬局）でのフィールドテストを行った。その結果、特にメタボリックシンドロームに関連した相関ルールが得られている[6]. 共同研究先から得られたフィードバックとして、これらのシステムでは実際の保健指導は従来どおり面接や電話、メールなどを使用するため、さらに具体的に保健指導に利用できるシステムが求められた[7].

図1は健康管理ASPシステムの概要である。システムのユーザ登録と初期設定は携帯電話から行う。このとき、ユーザはシステムで収集する項目を選択する、また、システムに用意しているデータ項目以外のデータをユーザが収集したい場合は、ユーザ自身が項目を追加する。実際の初期設定では、システムを利用するユーザの作業を簡便化するため、ユーザはあらかじめ用意されたエディションを選択する。ユーザが選択で

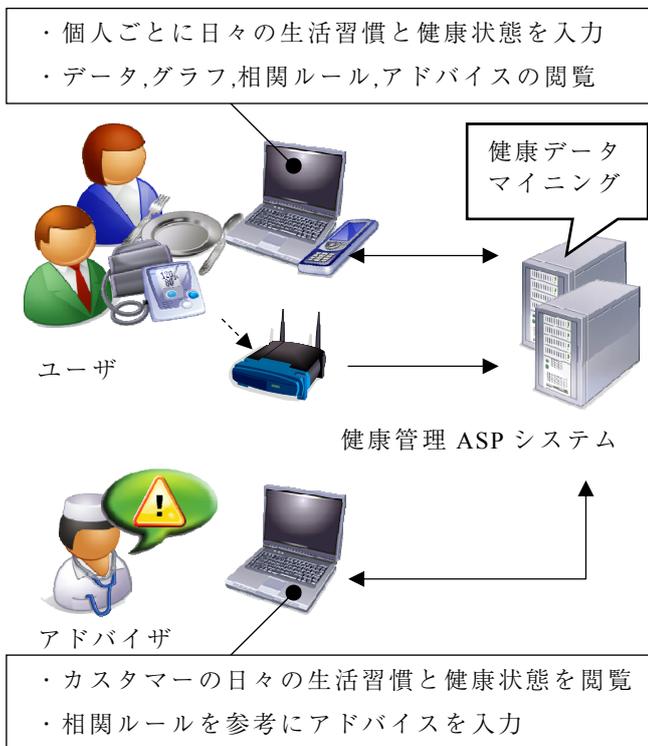


図1 健康管理ASPシステムの概要
fig.1 Outline of ASP system for healthcare

きるエディションはダイエット、ライト、プロ、ベットの、4種類で、例えばダイエットエディションでは生活習慣として摂取・消費カロリー、健康状態として体重、体脂肪率、ウエストサイズが項目登録されるなど、各エディションに応じて必要と思われる項目、追加項目がセットされる。システムの利用登録が終わると、ユーザは携帯電話またはPCのブラウザを用いて、インターネット経由で自身の日々の生活習慣と健康状態のデータをシステムに入力する。また、一部の健康データ（体重、血圧）については、BTによりホームゲートウェイと無線接続することでのDBへの計測値の自動入力に対応している。なお、BT機器による自動入力に関しては、準備が可能な機器が1セットであったため、ユーザから利用希望者を募った。

健康管理ASPシステムでは、健康データマイニングサーバが3ヶ月ごとに登録されたデータに基づいた個人ごとの生活習慣と健康状態に関する相関ルールを導出する。アドバイザは、各ユーザの日々の記録と相関ルールを参考に、PCよりインターネット経由でアドバイスを登録する。ユーザの画面には、入力された自身のデータやグラフ、抽出された相関ルール、アドバイザからのメッセージが表示され、ユーザはこれらの情報を閲覧して個人の健康管理に役立てる。

2.2. 健康データマイニング

我々が開発中の健康データマイニングでは、生活習慣と健康状態の関係を「生活習慣の蓄積が健康状態に

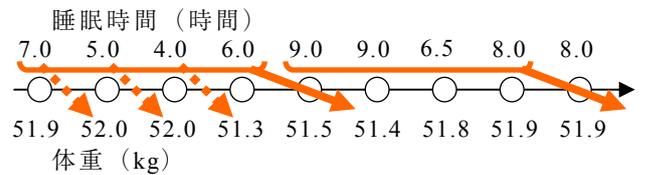


図2 時系列データ処理の模式的説明図
fig.2 Diagram of time-series-data processing

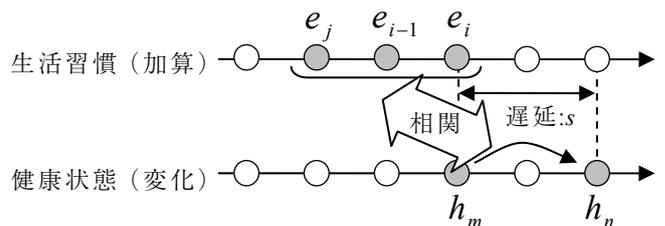


図3 時系列データ処理の概念図
fig.3 Concept of time-series-data processing

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」というシンプルな概念でモデル化している(図2)。つまり、ある健康状態(体重や体脂肪率、血圧、体温など)の変化を目的変数とし、その前日から数日前までに起こった生活習慣(摂取・消費カロリー、睡眠時間、運動量など)の蓄積を説明変数として、個人ごとの相関ルールを抽出する。

健康データマイニングのアルゴリズムでは、データのスクリーニング、相関ルールの生成と評価というスキームで相関ルールを抽出する。まずデータの事前処理として、データ数のチェックを行う。このとき、全体の標本数に問題がないデータ項目であっても、周期的な欠損があると時系列移動加算においてのデータのエントリ数が極端に減り、異常に少ないデータ数で相関係数を計算することになるため、あらかじめ欠損データの補完を行う[8]。次に、注目する健康データ項目とある程度関係がありそうな生活習慣データ項目を事前にスクリーニングするため、式(1)で表されるピアソンの積率相関係数を

$$r(\Delta h_{nm}, e_{ij}^t) = \frac{\text{Cov}(\Delta h_{nm}, e_{ij}^t)}{S.D.(\Delta h_{nm})S.D.(e_{ij}^t)} \quad (1)$$

計算する。図3の例では、1～数日前からのある健康状態(h)の変化と、その前の1～数日分のある生活習慣(e)の蓄積との相関係数を求め、この $r(\Delta h_{nm}, e_{ij}^t)$ が閾値を超えたものをデータマイニングの入力変数の候補に加える。

次に、スクリーニングで採用された生活習慣の蓄積を入力変数 Y とし、対象とする健康状態をその時系列データが「高い」「中間」「低い」の3つのシンボル値を持つ出力変数 X として、ITRULEアルゴリズム[9]を用いた相関ルールマイニングを行う[10]。ITRULEアルゴリズムは、

$$\text{If } Y=y, \text{ then } X=x \text{ with probability } p \quad (2)$$

という相関ルールを生成する。ルールの評価には、情報量の大きさの目安である J 測度[9]を用いる。

2.3. 相関ルールの抽出

前述の健康管理ASPシステムを用いて、16名のボランティアユーザーについて8名が①群として期間a(2010

年6月1日～8月31日)と期間b(9月1日～11月30日)、他の8名が②群として期間a(2011年6月1日～8月31日)と期間b(9月1日～11月30日)においてユーザー自身が興味のある生活習慣と健康状態のデータを日々蓄積し、健康データマイニングによって相関ルールの抽出を試みた。なお、性別は①群が男性5名、女性3名、②群が男性1名、女性7名である。

3. 結果

3.1. 健康管理ASPシステムの利用端末

ボランティアユーザーの端末選択について、①群では4名のユーザーが主にPCを用いてデータを収集し、4名のユーザーが主に携帯電話を用いた。②群では主にPCを用いたユーザーが3名、主に携帯電話を利用したユーザーが4名であった。また、ユーザーからは主に使用する端末ではない端末を使用してデータを入力した場合に誤入力が増えるという意見があった。BT機器による自動入力に関しては、現在では体重計と血圧計が対応しており、機器は1セットを準備している。このため①群、②群ともに1名が利用可能であったが、これを希望したユーザーは①群の1名のみで、②群での利用希望はなかった。

3.2. 相関ルール

健康データマイニングによって16名のボランティア

表1 ユーザ別の相関ルール抽出数
table 1 Number of association rules extracted by user

User	①a	①b	②a	②b
A	4			
B	1			
C	7	8		
D	12	7		
E	6	5		
F	14	9		
G		8		
H			7	3
I			4	
J			4	6
K			9	7
L			5	4
M				5
合計	44	37	29	25
平均	7.3	7.4	5.8	5
抽出人数	6	5	5	5
抽出率	0.75	0.63	0.63	0.63

アユーザについて抽出された相関ルール数の集計を表 1 に示す。相関ルールが抽出されたユーザ数の割合は概ね 6~7 割であった。

ユーザごとに抽出された相関ルールのサンプルを示す。

- A) 10 日間の自転車による運動の平均速度が 1 日平均 6.2km/h より速く、かつ 10 日間の自転車による運動時間が 1 日平均 10.1 分より多いならば、3 日後の体温が低い
- B) 2 日間の入浴時間が 1 日平均 16.2 分より長いならば、3 日後の最大血圧が高い
- C) 青汁摂取量が 1.5g 未満ならば、3 日後の内臓脂肪が高い
- D) 7 日間のヨーグルト摂取量が 1 日平均 153.6 g 未満ならば、3 日後の内臓脂肪が高い
- E) 4 日間の野菜ジュース摂取量が 1 日平均 237.5ml より多いならば、3 日後の肌水分が低い
- F) 2 日間のサプリメント (Ca&Mg) 摂取量が 1 日平均 2.2 粒より多く、かつ 2 日間の総摂取カロリーが 1 日平均 2212.2kcal より多いならば、2 日後の骨量が高い
- G) 睡眠時間が 6.2 時間未満で、かつ 7 日間のサプリメント (コラーゲン) 摂取量が 1 日平均 3.6g より多いならば、翌日の体重が高い
- H) 3 日間の総摂取カロリーが 1 日平均 1621.7kcal より多いならば、翌日の体重が高い
- I) 9 日間の総摂取カロリーが 1 日平均 1772.4kcal より多いならば、翌日の体脂肪率が高い
- J) 総摂取カロリーが 1300.5kcal 未満で、かつ 10 日間の筋トレ時間が 1 日平均 17.9 分より多いならば、3 日後の体脂肪率が低い
- K) 9 日間の総摂取カロリーが 1 日平均 1614.6kcal より多く、かつ 10 日間のりんご酢摂取量が 1 日平均 21.5ml 未満ならば、翌日の肌水分が低い
- L) 3 日間の睡眠時間が 1 日平均 6.9 時間未満ならば、2 日後の骨量が高い
- M) 6 日間の総消費カロリーが 1 日平均 310.5kcal 未満で、かつ 10 日間の総摂取カロリーが 1 日平均 1509.0kcal 未満ならば、3 日後の体脂肪率が高い

4. 考察

4.1. 利用端末について

ユーザが端末に携帯電話を用いるか PC を用いるかについてはユーザの任意としたが、①群、②群ともに主に PC と主に携帯電話を用いたユーザがそれぞれみられた。これはユーザの端末における習熟度が現れているのではないかと考えている。また誤入力に関して、誤入力のあったユーザから使い慣れない端末を使用したときに多くなるという意見があったのは、PC と携帯電話で UI が大きく異なることに起因するのではないかと考えられる。計測値の誤入力を防ぐためにはデータの自動入力があるが、BT による自動入力について利用者希望は 1 名に留まった。これは、現在対応している機器の接続や設定には家庭内 LAN 設備や BT のペアリングなど、IT に関するリテラシーが必要であることが原因し、使用の際の閾が高くなっているのではないかと考えている。

4.2. 相関ルールと個人健康管理

相関ルールが抽出されたユーザについては、各ユーザが自身の生活習慣と健康状態に関して何らかの洞察を得ている。

相関ルールのサンプルを挙げたユーザにおいては、例えばユーザ A では体温 (平熱) が低いということに関心のターゲットとしていた。このユーザは運動によって平熱の状態が改善するのではないかと考え、自転車による運動と体温についての関連性を知ろうと試みた。a の期間で抽出された相関ルールは運動を行うと体温が低くなるという結果となっているが、これは気温の高い夏季に抽出された相関ルールであるためだと考えられた。このユーザに関して b の期間 (秋季) では相関ルールは抽出されなかったが、追加解析の結果この期間では運動と体温に負の相関がみられた。このことから、運動によって体力が増強され、身体の恒常性が高まったのではないかと推察された。体温の恒常性について、体力には行動体力と防衛体力があり、防衛体力には暑熱・寒冷に対する抵抗力が含まれる。ユーザ A の場合は、自転車において 10 日間の運動量が 1 日に平均速度で 6km/h 以上、時間にして 10 分以上になるように継続すると効果が現れるのではないかとという洞察が得られた。

ユーザ H における関心事は体重の減少であり、主に

食事との関連を知りたいというものであった。このユーザでは 1 日の摂取カロリーの総量が 1 日平均 1,600kcal 程度を超えると、体重が増加するというルールが得られた。このユーザは 20 代女性であるが、20 代女性の基礎代謝量は概ね 1,200kcal といわれていることから、このユーザの生活習慣下では基礎代謝を 400kcal 程度超えてカロリー摂取が行われると体重の増加として現れるのではないかという洞察が得られた。すなわち、このユーザが体重を減らすためには、摂取カロリーを基礎代謝 + 400kcal 以下に抑えるか、400kcal に相当する日常的な運動を取り入れることが有効ではないかと考えられる。ユーザ K においても関心事は体重の減少であり、りんご酢の摂取と関連があるか知りたいというものであったが、このユーザでは体重との相関はみられず、主に肌水分との相関があることがわかった。いずれのユーザにおいても、生活習慣がどのような健康状態に影響しているかという、自身の健康管理に有用な洞察を得ることができた。従来の疫学研究における介入研究では、介入郡と非介入郡の分布の差から介入に効果があるかを探る試みといえるが、介入による感受性は個人によってばらつきがある。このことから、個人について時系列に蓄積されたデータから、その個人ごとの相関ルールを抽出することで、ユーザの健康管理において役立つ情報が得られたのではないかと考えている。

①群、②群ともに a, b の期間それぞれで抽出された相関ルールのユーザ比率が 6~7 割であったのは、ユーザが自主的に選択した生活習慣と健康状態のデータ項目に相関関係が見られない場合もあるためである。しかし、ユーザが自主的に関連しているだろうと考えている項目に相関が見出せない場合も、ルールが抽出されなかったということがユーザの健康管理における洞察となるため、有用な知識となり得る。

4.3. 課題

健康管理 ASP システムでは、PC や携帯電話を端末として使用できるが、今回の試みではユーザの利用端末がなぜ PC と携帯電話に分かれるのかを判断するには至らなかった。また、普段使用する端末以外を使用した場合に誤入力が増加することについての対策も必要である。

健康データマイニングによる相関ルールの抽出で

は、ルールが抽出されないユーザについてのモチベーションの維持に目を向けなければならない。本システムではユーザに 3 ヶ月間のデータ蓄積を要求しているが、これだけでもユーザにとって負荷となっている。この期間データを蓄積しながら相関ルールが抽出されなければ、ユーザはシステムの利用を継続できずに脱落する可能性が考えられる。

5. まとめ

健康管理 ASP システムにおいては利用端末として PC と携帯電話に対応しているが、ユーザが選択する端末は習熟度やリテラシーに依存するのではないかとということが示唆された。データの誤入力や自動入力機器におけるユーザのリテラシー不足など、システムとユーザの間に存在する課題も見えてきた。

相関ルールの抽出についてはユーザが興味のある生活習慣と健康状態を入力し、6~7 割のユーザについて相関ルールが抽出された。ユーザのなかには相関ルールが抽出されなかったり、思うようなルールが抽出されなかったものもあったが、総じては生活習慣と健康状態に関する何かしらの洞察が得られたものと考えている。今後はルールが抽出される、されないに関わらずユーザが何かしらの洞察を得られるような工夫を行っていく必要がある。

謝 辞

本研究は文部科学省科研費（課題番号：23500813）の助成を受けている。また、日本データベース学会と日立製作所による日立 HiRDB アカデミック制度の適用を受けている。

文 献

- [1] 厚生労働省大臣官房統計情報部, "平成 21 年度国民医療費", 2011.
- [2] H. Kumpusch, D. Hayn, K. Kreiner, M. Falgenhauer, J. Mor, and, G. Schreier, "A Mobile Phone Based Telemonitoring Concept for the Simultaneous Acquisition of Biosignals and Physiological Parameters", Proc. 13th World Congress on Medical and Health Informatics (Medinfo2010), pp.1344-1348, 2010.
- [3] H. Takeuchi, N. Kodama, T. Hashiguchi, and N.

Mitsui, "Healthcare data mining based on a personal dynamic healthcare system," Proc. 2nd Int. Conf. on Computational Intelligence in Medicine and Healthcare, pp.37-43, Lisbon, Portugal, Jun. 2005.

- [4] 竹内裕之, 児玉直樹, 橋口猛志, 林 同文, "個人健康管理を目的とした健康データマイニングシステム" 第17回データ工学ワークショップ論文集, 1B-i11, Mar. 2006.
- [5] 黛勇氣, 竹内裕之, 児玉直樹, 佐藤恵一, "データマイニング機能を備えた ASP システムの開発", 医療情報学, vol.30, no.3, pp.165-171, Jun. 2011.
- [6] 上村勝美, 黛勇氣, 児玉直樹, 竹内裕之, "特定健診・保健指導を支援するシステムの開発", 第19回データ工学ワークショップ論文集, C5-4 pp.1-6, Mar. 2008.
- [7] 黛勇氣, 竹内裕之, 児玉直樹, 矢部 貴之, "健康データマイニングを活用した特定保健指導支援システム", 第1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, E2-2 pp.1-5, Mar. 2009.
- [8] 竹内裕之, 黛勇氣, 児玉直樹, "健康データマイニングにおける欠損データの補完", 第2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, B8-2 pp.1-4, Mar. 2010.
- [9] P. Smyth and R. M. Goodman, An information theoretical approach to rule induction from databases, IEEE Trans. Knowledge and Data Engineering, vol.4, no.4, pp.301-316, Aug. 1992.
- [10] 竹内裕之, 児玉直樹, 橋口猛志, 林 同文, "個人健康管理システムのための自動相関ルール抽出アルゴリズム", DBSJ Letters, vol.5, no.1, pp.25-28, Jun. 2006.