情報処理学会 Interaction '97 (1997/02/04 ~ 05 於 東大山上会館)

個人適応型WWWのためのユーザモデル構築

三浦 信幸　　高橋 克巳　　島 健一

NTT ソフトウェア研究所

E-mail : {miura, takahasi, kshima}@slab.ntt.co.jp

提供する情報内容および情報提示方式をサーバ自身がユーザ個々に適応していく個人適応型 WWW を構築するために重要なユーザモデルについて検討を行った．本稿では，まず，我々の目指す個人適応型 WWW を定め，個人適応型 WWW の構成を提案する．この中で重要な構成要素のひとつであるユーザモデリングに関して，ユーザモデルのモデルの枠組，ユーザモデル構築に必要な WWW コンテンツモデルの概要，ユーザモデル構築方法などを検討した．モデル表現をグラフ構造とし，モデル構築方法をそれらのグラフに対するグラフ操作としていることがこれらのモデル・構築方法の特徴である．さらに，検討結果に沿って，実際のサーバ・ログから，ユーザモデルを構築する実験を行い，従来，想定されているようなユーザの単純な特性よりも，より多様な，より興味深いユーザの特性を抽出できることを確認し，検討したユーザモデルの有効性を示した．

キーワード:　WWW，ユーザモデル，個人適応型 WWW，パーソナライズド Web

A User-Model Construction Method for a Personal-Adapted WWW

Nobuyuki Miura　　Katsumi Takahashi　　Ken'ichi Shima

NTT Software Laboratories

This paper describes a user-model construction method for a personal-adapted WWW. In a personal-adapted WWW, contents of information and how to display them are adapted to users' preferences. At first, this paper introduces features and an architecture of our personal-adapted WWW. Especially, this paper focuses user-modeling which is one of the most important parts of this architecture. This paper proposes frameworks of user-models, WWW contents-models and a method for constructing user-models. Our method has two characteristics that models are represented as graphs and the construction method is based on the graph-operation. And we show experimental results of constructing user-models from real WWW server logs. The results show effectiveness of our method in the viewpoints that we can get more various and more interesting users' profiles than what are produced by usual construction methods.

Keywords:　WWW, User Models, Personal-Adapted WWW, Personalized Web

1 はじめに

昨今の WWW の広がりにつれて，より質の高い，より高度で知的な WWW が求められる時代となってきた．本稿では，そのような要求に答える WWW として，個人適応型 WWW を提案する．従来の一般的な WWW が基本的には，全てのユーザに対して，同一の情報内容を同一の情報提示方式で提供してきたのに対し，個人適応型 WWW は，提供する情報内容・情報提示方式等をユーザ一人一人の特性に応じてユーザ毎に動的に変更し，ユーザに適応しながら動作する WWW である．

現在，いくつかのサイトにおいて，あらかじめ申告しておいたユーザ特性に基づき，サーバが必要なコンテンツを引出し，画面整形などを行って情報提供するものがある (例えば，[1, 2])．また，このような個人対応を行うためのソフトウェア・パッケージも出始めている[1]．このような個人対応を行う WWW はパーソナライズド Web とも呼ばれ，今後の Web の方向性のひとつとして，現在，非常に注目されている分野である．

しかしながら，現在のパーソナライズド Web のほとんどは，使用開始時に自分のユーザ特性をあらかじめ申告しておき，それをもとに個人対応を行うだけで，使用開始以降のユーザの変化に追従していかないものが多い．また，行われる個人対応についても，ユーザ特性とコンテンツ内の各 URL とをキーワードで対応付けて情報フィルタリングする，画面レイアウトの単純な変更を行う，などといった，比較的，単純な仕組みのものが多い．より高度な個人適応型 WWW 実現のためには，ユーザ特性の把握・解釈，それを用いたコンテンツの個人適応の仕組みを根本的に深く考え直す必要性がある．

本稿では，まず，2章で我々の目指す個人適応型 WWW の特徴，個人適応の例を紹介する．これに基づき，3章で個人適応型 WWW の構成を提案する．本稿ではこの構成の中で特に重要となるユーザモデルの構築に関して検討を行った．まず，4章でユーザモデルの枠組を提案し，5章でユーザモデル構築に必要な WWW コンテンツモデルの一例を紹介し，6章ではユーザモデルの構築方法を提案する．さらに，7章で，提案手法の有効性を確認するためのユーザモデル構築実験の結果の一部を紹介する．最後に 8章でまとめと今後の課題について述べる．

2 本研究が目指す 個人適応型 WWW

現在，個人対応を実現しているいくつかの WWW サイトを参考にして，我々が目指す個人適応型 WWW には，次のような特徴を持たせることを目標としている．

- ユーザは通常通りアクセスするだけで適応
  従来の個人対応 WWW サイトのほとんどは，事前に，キーワード群の中から自分の興味等を選択してアンケートとして提出する方式である．我々が目指す個人適応型 WWW では，サーバ側でユーザの特性を自動的に取得し，サーバ自身が個人に適応するため，そのようなアンケートは基本的には不要である．
- アクセスの繰返しによるユーザの変化に追従
  事前アンケート形式では，特にユーザが自主的にアンケート内容を変更しない限り，個人対応形式は変化しない．我々が目指す個人適応型 WWW では，サーバは常にユーザの最新の特性に合わせて個人適応を行う．
- ユーザの行動パターンや学習状況等にも適応
  アンケート形式では，ユーザの興味・関心等のいわば，ある特定の瞬間のユーザの状態のみにしか対応ができない．我々が目指す個人適応型 WWW では，ユーザの行動パターンやユーザの該当 WWW のコンテンツ・機能に関する学習状況の特性についても個人適応を行う．

次に，個人適応型 WWW で行う，個人適応の例を現在考えられる範囲内でいくつか紹介する．題材として，我々が現在試験提供している英語 TOWNPAGE[2] およびインターネットタウンページ[3] を想定してみることにする．これらの WWW では，職業別電話帳の内容を検索することができるとともに，広告情報，地図情報，各種観光情報・流行情報などが統合的に提供されている．

- 業種名検索における業種名リスト並び変え
  業種名から電話帳掲載を検索する場合に，業種名のリストの中から業種名を選んで検索を行うインタフェースとなっている．ユーザの検索履歴を利用して，そのユーザにとって関連があると思われる業種をリストの先頭の方に表示するように業種名リストを並び変える．
- ユーザ特有の関連業種を表示
  ある業種について検索を行った後に，関連する業種一覧を表示する機能がある．この一覧は，一般的に関連があると思われる業種のリストである．これに加えて，ユーザの検索履歴の情報を基に，そのユーザ独自の関連業種を優先的にリストアップする．
- 行動パターンと ユーザの興味を考慮した広告表示
  多くのユーザが，引越し業者の検索の後に，家具屋の検索を行っていたとする．また，あるユーザは輸入物に関する検索を良く行っていたとする．このような状況において，このユーザが引越し業者の検索を行った際に，輸入家具屋の広告を表示する．

[1] http://www.firefly.net/，
http://www.compassware.com/等

[2] http://www.pearnet.org/jtd/

[3] http://townpage.isp.ntt.co.jp/

- 行動パターンに従った検索結果の絞り込み
  多くのユーザが，レストランの検索の後にホテルの検索を行っていたとする．この時，あるユーザがレストランの検索の後にホテルの検索を行った場合に，前回検索を行ったレストランにより近いホテルを優先的に表示する．
- ユーザのコンテンツに関する理解度に応じた機能ガイド
  あるユーザが，このWWW特有の機能である検索結果を地図で表示できる機能を全く使っておらず，この機能の存在を知らないと予想されるとする．地図機能が有効に機能すると考えられる場面 (例えば，前述のレストランの近くのホテルを優先的に表示している場面) にこのユーザがたどり着いた時，地図機能の存在をアピールする文面等を画面に追加する．

このような個人適応をheuristicに行うのではなく，体系的に行えるようにするための個人適応型WWWの構成法およびユーザモデリングを検討する．

3 個人適応型 WWW の構成

WWWの世界では，サーバのコンテンツ提供者とコンテンツをブラウズする利用者が存在する．両者はそれぞれ，一般に複数人であり，各個人はそれぞれ異なる背景知識や興味をもっている．つまり，複数人のコンテンツ提供者と複数人の利用者の間には，提供されているコンテンツの認識に関して，何らかのギャップが存在している可能性がある．そのギャップを埋めることこそが，個人適応であると考えることができる．したがって，まず，コンテンツ利用者のモデル (ユーザモデル) とコンテンツ提供者のモデル (コンテンツモデル) を作成することが必要であり，この両者のモデルを擦り合わせることによって，どのように個人適応するか，すなわち，コンテンツをどのように加工してユーザに提示するかのルールを作成することができると考えられる．特定のユーザモデルとコンテンツモデルを擦り合わせてその個人用のルールを生成するほかに，複数のユーザのユーザモデルどうしの擦り合わせによるルール生成も考えられる．後者は，[3] などで採用されているSocial Filtering，あるいは，Collaboration Filteringと呼ばれる考え方を利用したもので，他者の履歴を利用してコンテンツを特定のユーザ用に適応させる (情報filteringする) 方法である．ここでは，後者の方法と区別するために前者の方法をPersonal Filteringと呼ぶことにする．

ユーザモデルの構築に必要なユーザ情報の取得法には，次のような方法が考えられる．

- ユーザにアンケートを取る．
- ユーザにヒアリングやインタビューを行う．
- ユーザの挙動を観察する．
  - クライアント側 (ブラウザ側) で観察する．
  - サーバ側で観察する．
  - クライアントとサーバの間の proxy server 上で観察する．

ここでは，なるべくユーザに余計な作業・コスト等の負担をかけないこと，つまり，基本的にはWWWサーバ側の作業のみで情報の抽出が可能なこと，また，ユーザの時間変化に追従できる情報が取れること，などの要件から，ユーザの挙動をサーバ側で観察する方法を採用する．WWWの場合，サーバ側で観察できるユーザの挙動のほとんどは，WWWサーバのログという形で保存されており，これを利用することにする．サーバ・ログを使ったユーザモデル構築では，個々のログについてユーザを特定することが困難 [4] という問題があったが，本研究では，HTTP Cookie [4] という仕組みを利用して，この問題を解決することとした．

クライアント側でユーザ情報を取得し，クライアント側で個人適応を行う方法も考えられるが，ユーザモデルの構築およびWWWコンテンツの個人適応にあたっては，コンテンツの内容・構成・機能といった情報が不可欠であり，それらの最新の情報が豊富に蓄積されているサーバ側でユーザモデルの構築や個人適応を行う方が効果的であると考えられる．

また，ユーザモデルの構成要素は，コンテンツモデルとの擦り合わせを行うために，コンテンツモデルの構成要素と類似のものである必要がある．したがって，ユーザモデルは，サーバ・ログとコンテンツモデルの両方を利用して構築される必要がある．

一方，コンテンツモデルは，コンテンツから得られる形式的な情報とコンテンツ製作者の製作意図を組み合わせてモデルを構築することとする．

以上を合わせると，個人適応型WWWの全体像は，図1のように表せ，個人適応型WWWの重要な要素として，ユーザモデリング，コンテンツモデリング，個人適応ルール生成 (Social Filtering, Personal Filtering) の3つがあげられる．この3つのうち，本稿では，特に，ユーザモデリングに関して述べる．

4 ユーザモデルの枠組

さまざまな研究分野において，人間のモデルを作ろうとする試みが古くから行われているが，ここでは，次のような観点からユーザモデルを考える．

- 個人適応型WWWのために必要なモデルであること．WWWサーバがユーザに提示する情報の内容や提示方式を個人毎に動的に変更するために必要な程度の情報量で良い．
- WWWサーバ・ログから取得可能あるいは類推可能な情報で構築可能なモデルであること．

2章で考えた個人適応の例に基づき，前者の観点から，次のようなユーザに関する情報が必要と考えられる．

[4] http://home.netscape.com/newsref/std/cookie_spec.html

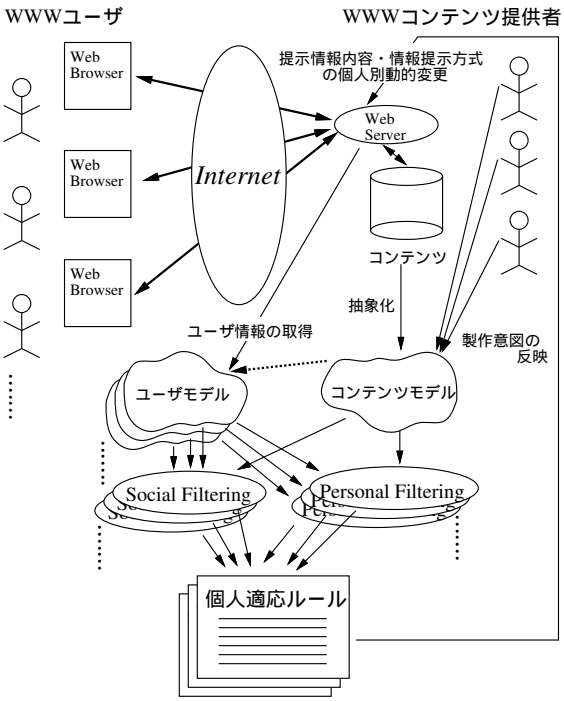

図 1: 個人適応型 WWW の全体像

- ユーザの興味・関心等のユーザの状態の情報
- ユーザの該当 WWW における行動パターン等のユーザの行動に関する情報
- 該当 WWW の構造・提供している情報の種類・機能等をユーザがどの程度，理解しているかの情報
- ユーザがどのような時にどのような行動を取り，どのようなことを学習したかといったエピソード的な情報

後者の観点から考えると，上にあげた情報は WWW サーバ・ログから，取得可能あるいは類推可能な情報である．

したがってこれらは，前者の要件，後者の要件ともに満たしているので，ここでは，この 4 つを個人適応型 WWW のためのユーザモデルとする．このようなモデルが適切かどうかの評価は，7章で述べる実ユーザのアクセス履歴からユーザモデルを構築する実験において，ユーザ毎の違いを表現できているか否かをもって判断する予定である．

これらそれぞれに合わせて，図 2のような状態モデル・行動モデル・学習モデルの 3 つのモデルとこれらのモデル間の関係を表現するメタモデルの計 4 つのモデルを構築する．個々のモデルの内容はおよそ次のようなものである．

- 状態モデル
  ユーザの興味・関心等を表現するモデル．
  具体的には，各 URL[5]毎に，その URL が表現している内容についてのキーワードを 1 個以上設定し，各 URL へのアクセス 1 回を各キーワードへのアクセス 1 回と抽象化してモデルを作る．
- 行動モデル
  ユーザの該当 WWW における行動パターン等を表現するモデル．
  具体的には，各ユーザが行ったハイパーリンクの遷移列のうち，頻出する遷移列をユーザの行動パターンとして抽出した遷移列の集まりである．この遷移列には，どれくらいの時間間隔で行動を起こしたかの平均値なども属性値として付く．
- 学習モデル
  該当 WWW に対するユーザの理解度を表現するモデル．
  具体的には，該当 WWW が持つ特徴的な構造・機能，提供している情報の種類などの該当 WWW にアクセスする上で学習していくと考えられる項目と，それらを使っているか否か・理解しているか否かとを対応づけたものである．個々の学習項目どうしは関連性のある項目どうしは互いにリンクを持つ．なお，各項目には，それぞれ，いつ学習が行われたかの時間も記録される．
- メタモデル
  状態・行動・学習の 3 つのモデル間の関係，特に時間的な同期関係を表現するモデル．
  具体的には，いつ，どういう状態モデルの時に，どのような行動パターンが発生し，どのような学習が行われたかの記録の集まりである．

[5] ここでは CGI 等の引数の query string を考慮に入れ，query string を含めた URL それぞれについてキーワードを設定する

図 2: 4 つのユーザモデルの関係

5 WWW コンテンツモデル

3章で述べたように，個人適応型 WWW には，WWW コンテンツモデルが必要である．このモデルは，ユーザモデルとの擦り合わせにより，どのような個人適応を行うかという判断の際に用いられるばかりではなく，図 1に書いたようにユーザモデル構築の際にも必要である．例えば，職業別電話帳の WWW では，業種別検索の URL に対して業種のキーワードが付与され，これを使って業種別検索に関するユーザモデルを構築する．

WWW のコンテンツはハイパーリンクという構造を既に持っていること，コンテンツ・モデルは一般に，ある程度複雑な構造を持つことが予想されること，の 2 点から考えて，コンテンツモデルは，グラフで表現しておく必要があると考えられる．例えば，2章でも題材にした，職業別電話帳の WWW の場合を考えると，図 3のようなグラフが

考えられる．これは，業種名をノードとして，関連のある業種をリンクしたものである．

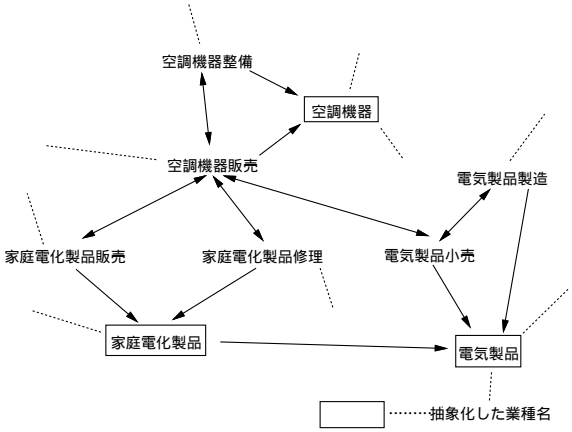

図 3: 電話帳 Web の職業別検索の
コンテンツモデルの例

コンテンツモデルの作成にあたっては，WWW コンテンツが一般にかなり複雑であること，コンテンツがどのようなドメインに属するものかによって，モデリング手法が異なると考えられること，3章でも述べたように，コンテンツ製作者の意図の反映という複雑なプロセスが必要であることなどから，モデリング手法の一般化は非常に難しい．当面は，各 WWW サイト毎に検討し，モデル作成を行うことになると考えられる．

6 ユーザモデル構築方法

3章で検討したように，ユーザモデルの構築にあたっては，ユーザ毎に分割したサーバ・ログと該当 WWW のコンテンツモデルが必要である．5章で述べたように，コンテンツモデルはグラフ表現しておく必要性があることもわかった．一方，サーバ・ログに関しても，図 4の例のようにグラフで表現することができる．図 4は，アクセス・ログの職業別検索の部分のみを取り出したもので，検索した業種名をノードとし，アクセス順にリンクをつけ，リンクにはアクセス間隔の時間が属性値としてついたものである．

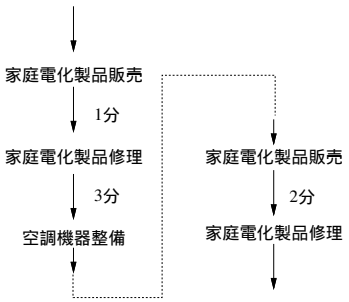

図 4: 業種別検索のアクセス・ログの例

このように，ユーザモデル構築に必要なコンテンツモデル，アクセス・ログともにグラフで表現することができるとわかった．また，ユーザモデルはこれら 2 つから構築されることから，ユーザモデルもグラフ表現することが適切であると考えられる．ゆえに，ユーザモデル構築方法としては，コンテンツモデル，アクセス・ログをグラフ表現したものに対し，必要なグラフ操作を施し，グラフ表現されたユーザモデルを構築する方法が妥当であると考えられる．構築方法の概要は図 5のように表せる．

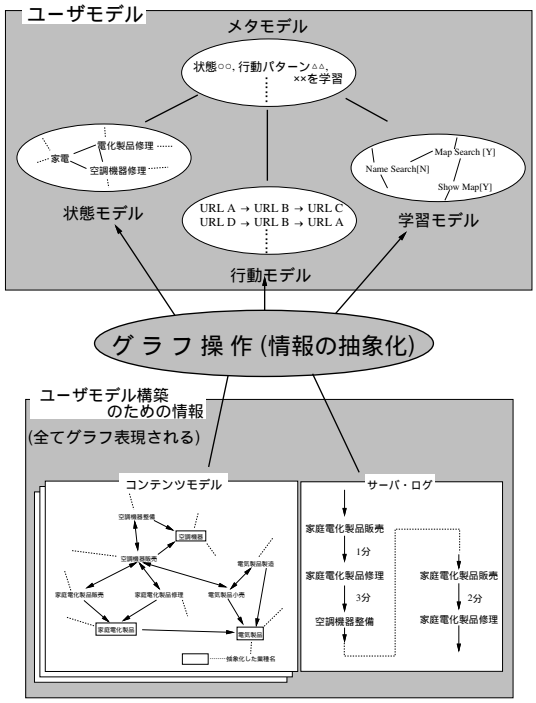

図 5: ユーザモデル構築方法の概要

では，具体的に，構築方法論を検討する．ここでは，電話帳 WWW の職業別検索に関する状態モデル作成を例として述べる．コンテンツモデルの例として図 3を，アクセスログの例として図 4を考える．この 2 つのグラフに対して，次のような手順で電話帳 WWW の職業別検索の状態モデルを作成する．

モデル構築法の基本的な考え方は，次のような仮定に基づいている．

- アクセス回数の多いノードは，そのユーザにとって興味や関心の高いノードである
  興味や関心の高いノードには，ユーザは何度もアクセスを行うと考えられる．
- たどった回数の多いリンクは，そのユーザにとって興味や関心の高いリンクである
  ユーザは，興味や関心の高いリンクを何度もたどると考えられる．
- 前後のノードにアクセスした時刻の差が短いほど，そのユーザにとって興味や関心の高いリンクである．

時間間隔が広ければ，そのユーザにとって前後のノード間の関連性は低いものと予想される．逆に，時間間隔が短ければ関連性は高いものと予想される．

これらの仮定に基づき，ノードに関してはアクセス回数の多い業種ノードはアクセス頻度が高いとし，リンクに関してはたどられた回数が多いリンクや短時間の間にたどられたリンクはアクセス頻度が高いとするようなモデル構築を行う．ユーザモデルを使って個人適応を行う際には，アクセス頻度が高いノードやリンクはユーザにとって興味・関心が高いものであるとモデルを解釈して，個人適応ルールを生成する．

また，ノード (業種名) のアクセス頻度とリンクのアクセス頻度との比較を容易に行うため，それぞれの頻度の値は $0 \sim 1$ の間を取るよう平均化する．

まず，コンテンツモデルとサーバ・ログのグラフを次のように定義する．

- コンテンツモデル
  ノード $a_i =$ 業種名 $(i = 0...n)$，ノード a_iから a_jへのリンクを $l_{i,j}$とする．また，ノード a_iのアクセス頻度を nf_i，リンク $l_{i,j}$のアクセス頻度を $lf_{i,j}$とする．
- サーバ・ログ
  ノード $b_i =$ 業種名 $(i = 0...k)$，ノード b_iから b_jまでの時間間隔を $t_{i,j}(i < j)$ とする．

これに従って，次のようにモデルを構築する．

1. サーバ・ログのグラフ上で同一業種のノード数をカウントし，その合計で平均化して個々の業種のアクセス頻度を計算する．これをコンテンツモデルの各ノードのアクセス頻度 nf_i とする．
   例えば，図 4では，家庭電化製品販売 2，家庭電化製品修理 2，空調機器整備 1 であるので，家庭電化製品販売 $2/(2+2+1) = 0.4$，家庭電化製品修理 $2/(2+2+1) = 0.4$，空調機器整備 $1/(2+2+1) = 0.2$ となる．
2. コンテンツ・モデルのグラフ上で，抽象化業種に関しては，子ノードのカウント数の平均値を抽象化業種名のノードに付与する．このことにより，抽象化業種も実在する業種も同じアクセス頻度という尺度で比較できる．
   例えば，図 3では，家庭電化製品販売 0.4，家庭電化製品修理 0.4 から，$(0.4+0.4)/2 = 0.4$ で家庭電化製品のアクセス頻度は 0.4 となる．
3. リンクのアクセス頻度 $lf_{i,j}$は，$t_{i,j}$の最小値を t_{min}とすると，

$$lf_{i,j} = \sum_{i=1,j=1}^{n,k} \begin{cases} if((a_i == b_i)\&\& \\ \qquad (a_j == b_j)) \\ \quad then \quad \dfrac{1/t_{i,j}}{1/t_{min}} \\ else \quad 0 \end{cases}$$

とする．$lf_{i,j}$は，そのリンクがたどられた回数が多いほど，また，たどられた際のアクセス間隔が短いほど，大きな値を取る．

4. アクセス・ログのグラフとコンテンツ・モデルのグラフの重ね合わせを行い，ユーザ毎のコンテンツ・モデルのグラフを作成する．重ね合わせた結果得られるグラフの構成要素は次のようになる．
   $nf_i > 0$ のノード，両端のノードが $nf_i > 0$ であるリンク．

以上の作業で抽出されたユーザの状態モデルは，図 6のようになる．

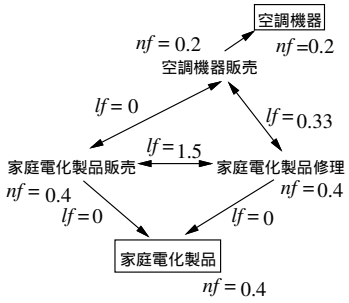

図 6: 抽出された状態モデルの例

このようにコンテンツ・モデルとアクセス・ログのグラフの重ね合わせにより得られるものは，一般的に記述されたコンテンツ・モデルの中から，各ユーザの特性に合った部分のみを抽出し，さらに各ユーザ特性を加えたものになる．図 6の例でいえば，得られたグラフはもとのコンテンツ・モデルのグラフの一部に相当するものであると同時に，家庭電化製品販売と家庭電化製品修理との間のリンクを加えたグラフである．新たにリンクを加えたことが，一般的なコンテンツ・モデルにユーザ特性を追加したことに相当する．

同様の手法で，電話帳の職業別検索以外の種類の URL に関しても，一般的に記述されたコンテンツ・モデルの中から，各ユーザの特性に合った部分のみを抽出し，さらに各ユーザ特性を加えたというユーザの状態モデルを作成することができる．

一方，行動モデルであるが，4章で述べたように，ユーザが行った頻出するハイパーリンクの遷移列をモデルとすることから，このような頻出遷移列の抽出作業がモデル構築法となる．そのような抽出アルゴリズムは，[5] で紹介されているものをはじめ，既に数多くのものが存在している．モデル構築実験を通じて検討する予定である．

さらに，学習モデルであるが，該当サイトのコンテンツに関して，ユーザが理解を深めていくと思われる，該当 WWW が持つ特徴的な構造・機能，提供している情報の種類などをコンテンツモデルとして表現しておく必要がある．また，それらの機能やページにどれくらいアクセスがあった際に，学習が行われたかということもモデルとして記述

しておく必要がある．これらのコンテンツモデルを利用して学習モデルは構築されるが，詳細は現在検討している最中である．

ユーザモデル構築にかけるコストと得られるbenefitの関係を考慮すべきであるという観点から，ユーザモデル全体の妥当性については，モデル単体で判断するのではなく，最終的に行った個人適応が妥当か否かをもって，判断したいと考えている．モデル単体で考えると厳密性に欠く部分もあるかも知れないが，行われる個人適応がそこそこ妥当であれば，個人適応型WWWのためのユーザモデルとしては妥当であると判断したい．

7 ユーザモデル構築実験

以下のような実験設定でユーザモデル構築実験を行っている．実験結果の情報量は膨大であり，詳細な分析はまだ作業中である．ここではその実験結果の一部を報告する．

- 実験対象WWW
  英語TOWNPAGE[6][6] ：
  本WWWは，英語版の業種別電話帳WWWであり，次の点から，本実験に適したWWWである．
  - 提供している情報が比較的明解に分類可能であり，客観的なコンテンツ・モデルが作成しやすい．従って，コンテンツモデルの出来・不出来がユーザモデル構築に影響する度合が低い
  - このWWWには，ユーザは通常，具体的な目的を持ってアクセスしてくるため，そこから作られるユーザモデルは個人適応を行うための適切なモデルになりうる．
  - 電話番号の掲載情報等の本来の情報の他に，広告や日本の観光案内といった副次的な情報も含まれており，多様なユーザの性質を把握するのに適している．
  - 1995年10月より実験提供しており，多数のユーザのログを利用することが可能である．
  - [4, 7]における予備実験において，確かにユーザの特性の把握ができている．
- 実験目的
  - 従来考えられているようなユーザモデリングに比べて，グラフ操作によるユーザモデリングが有効であることを確認する．
  - 作成されたモデルが確かにユーザ毎に違いのあるモデルであることを確認し，4章で検討したユーザモデルの枠組が妥当であることを確認する．
    また，モデルのうち，どの部分がユーザ毎の相違が大きいかを調べる．ユーザモデルの再検討材料とすると同時に，相違の大きい部分こそ個人適応すべき部分であるため，個人適応の具体的内容・具体的方法の検討材料とする．
- 実験方法
  モデル構築ツールのプロトタイプを実装し，それをサーバ・ログに適用して，実際にモデル構築を行う．

ここでは特に，職業別検索における職業分類に関するコンテンツモデルに着目し，職業分類に関するユーザの状態モデルのユーザモデル構築実験結果を考察する．コンテンツモデルは，約600のノードと約800のリンクを持つもので，電話帳発行元が長年のノウハウによって関連性のある業種をリンクしたものに，このWebを構築したコンテンツ提供者が抽象化業種を加えたモデルを採用した．

図7は，実験対象WWWのアクセス・ログから実際に構築した，あるユーザの職業分類に関する状態モデルである．このユーザは，1996年4月1日~6月20日にかけて，合計157回の電話番号検索を行っており，そのうち，職業名からの検索を87回行っていた．この87回分のデータから，このユーザの職業分類に関する状態モデルを6章で述べた方法で構築した．構築されたこのユーザの状態モデルには，コンテンツモデルから抽出された14ノード，26リンクと，ユーザのアクセス・ログから追加された8リンクが含まれている．

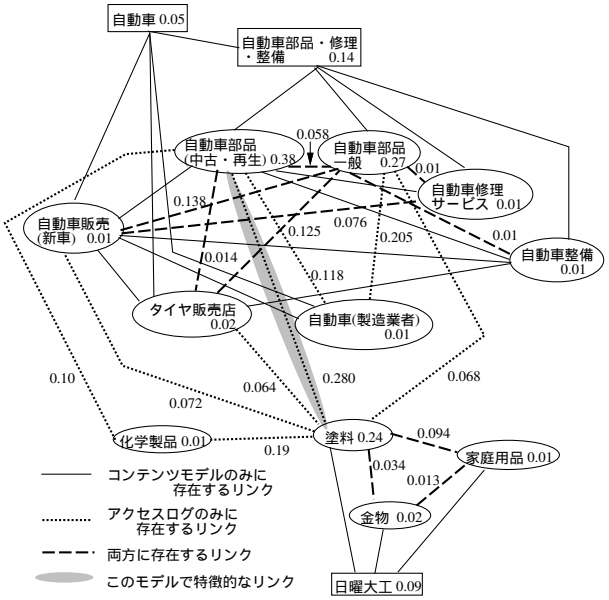

図 7: モデル構築実験結果の一例

一方，一般に想定される，単純なユーザモデル構築法では，表1のようなモデルが構築され，このユーザは，主に，自動車に関する業種と日曜大工に関する業種の検索が行っていると判断されることが多い．

しかし，今回提案する手法で得られたグラフ全体を見ると，塗料の業種を媒介にして，両者の業種は強い関係をもっている．特に，“塗料” と “自

[6] http://www.pearnet.org/jtd/

表 1: 従来のモデリング法によるユーザモデル例

自動車	71%	自動車部品・修理・整備	67%	自動車部品(中古・再生)	38%
				自動車部品一般	27%
				自動車修理サービス	1%
				自動車整備	1%
		その他	4%	タイヤ販売店	2%
				自動車販売(新車)	1%
				自動車(製造業者)	1%
日曜大工	27%			塗料	24%
				金物	2%
				家庭用品	1%
その他	1%			化学製品	1%

動車部品 (中古・再生)" との間のリンク (図 7中の影付のリンク) が頻繁にたどられている．これらは，このユーザにとっては，この 2 種類の業種が何らかの関係を持つものと考えられていると解釈できる．単純なユーザモデリングよりも，本稿で提案しているグラフ操作によるモデリングの方が，より多様なユーザの情報を表現できるということの典型的な例であると言える．

このように，本稿で提案するユーザモデル構築方法では，従来のノードのアクセス回数のみを基にした単純なユーザモデル構築法では決して把握することのできない，多様なユーザの特性を把握できることを確認し，提案するユーザモデル構築方法の有効性を示した．

このようにして，作成されたユーザモデルから個人適応の動的変更ルールを生成する際には，ユーザが現在参照している URL に相当するノードを始点として，ユーザモデル上のリンクやノードをたどって必要な情報を取得して，個人適応ルールを生成することになる．

8 まとめと今後の課題

本稿では，まず，個人適応型 WWW の全体像とそのシステム構成を提案した．これに基づき，重要な部分を占める，ユーザモデル構築について，ユーザモデルの枠組，ユーザモデル構築に必要な WWW コンテンツモデルの概要，ユーザモデル構築方法を検討した．モデルやサーバ・ログを全てグラフで表現し，それらに対するグラフ操作でモデルを構築することが本方法の特徴である．更に，実際のアクセス・ログからモデルを構築する実験を行い，従来の単純なユーザモデル構築方式よりも，より多様な，より興味深いユーザの特性を抽出できることを確認し，検討したユーザモデルの有効性を示した．

今後の課題としては以下のようなものを考えている．

- 行動モデル・学習モデル・メタモデルの詳細の検討
  本稿では主に状態モデルの枠組・構築方法・評価実験について論じた．残りの 3 つのモデルについても同様の検討を行い，ユーザモデルの全体を構築する．
- アクセス間隔を考慮したアクセス頻度の計算法
  6章で述べた方法では，コンテンツモデルのノードのアクセス頻度を単にアクセス回数で計算しているが，アクセス間隔が広い場合と狭い場合では，そのアクセス頻度は違う意味を持つものと考えられる．また，人間の記憶の忘却曲線 [8] の性質から考えても最後のアクセスから現在までの経過時間によってアクセス頻度の解釈が変わるものと考えられる．[9] で検討されているように，アクセスの時系列の傾向を考慮した，アクセス頻度の計算が必要であると考えられる．
- 個人適応以外の利用法
  現在は，動的に個人適応を行うことを考えているが，作成されたユーザモデル，特に行動モデルを利用することによって，各 page のユーザインタフェースを改善していくことも考えられる．類似研究，例えば [5] 等を参考にしたい．
- 他の WWW サーバへの適用の試み
  現在は，英語版 TOWNPAGE，あるいは，インターネットタウンページを適用例として検討しているが，これらとは異なるドメインの WWW の場合に，現在検討している手法がどこまで有効で，何を補わなければならないのかを検討したい．
- 教育学や心理学的見地からの個人適応法の検討

参考文献

[1] Internet TV Guide. http://www.tvguide.or.jp/.

[2] My Yahoo. http://my.yahoo.com/.

[3] W. Hill, L. Stead, M. Rosenstein, and G. Furnas. "Recommending and Evaluating choices in a virtual community of use". In *CHI '95*, pp. 194–201. ACM, 1995.

[4] 三浦信幸, 島健一. サーバアクセス履歴からのユーザモデルの構築. 情報処理学会 第 52 回 全国大会, Vol. 1, pp. 153–154, Mar. 1996.

[5] 岡田英彦, 旭敏之, 井関治. "使いやすさ評価ツール"GUI テスタ"における共通操作パタン抽出方式の提案と評価". 情処研報, Vol. 95, No. 104, pp. 37–42, Nov. 1995. 情報処理学会 研究報告 ヒューマンインターフェイス研究会 95-HI-63 No.7.

[6] 島健一, 高橋克巳, 三浦信幸. インターネット版マルチメディア電話帳の構築. In *Japan World Wide Web Conference '95*. 日本インターネット協会, Nov. 1995.
http://www.pearnet.org/jtd/paper/jtd-overview.html.

[7] 三浦信幸, 島健一. "WWW サーバ・ログからのユーザモデル構築実験". 情報処理学会 第 53 回 全国大会, Vol. 4, pp. 158–159, Sep. 1996.

[8] Ebbingahus H. von. *"ÜberdasGedächtnis"*. Duncker und Humboldt, 1885. 訳本「記憶について」誠信書房 (1978).

[9] 畑島隆, 元田敏浩. "WWW アクセスログの有効な解析法について". 情報処理学会 第 53 回 全国大会, Vol. 3, pp. 217–218, Sep. 1996.