

北海道観光情報ポータルサイトにおけるユーザモデルの構築

北海道大学大学院工学研究科
1004631 北海道大学大学院工学研究科* 大野貴司 OHNO Takashi
大内 東 OHUCHI Azuma

1. はじめに

北海道観光情報 WWW サイト数の急速な増加に伴い、ユーザが効果的に目的とする情報を収集することが可能なポータルサイト構築の必要性が高まっている。本稿では、個人適応型 WWW におけるユーザモデル構築法 [1] を用いて、北海道観光情報ポータルサイトの構築について考察する。

2. 個人適応型 WWW の基本的な概念とその構成

個人適応型 WWW の基本的な概念として、「ユーザになるべく余計な作業負担をかけずにユーザ特性を把握できること」、「ユーザ特性の時間変化に追従できること」、さらに、「ユーザの行動パターン等にも適応可能なこと」等が指摘されている [1]。本稿では、題材として「北海道観光情報ポータルサイト」を想定する。

WWW の構造は、大きく分けてコンテンツの提供者とその利用者から構成される。一般的に、両者は異なる背景知識や興味・関心を持つものと考えられる。つまり、提供されるコンテンツの認識に関して両者の間に何らかのギャップが存在する可能性がある。このギャップを埋めるためには、コンテンツ提供者のモデル（コンテンツモデル）とその利用者のモデル（アクセスモデル、ユーザモデル）を作成し、両者を比較・検討することが必要である。

3. ユーザモデルの構築法

まず、コンテンツモデルについて検討する。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や、コンテンツ同士の意味的な関係性を表現するために、コンテンツモデルをグラフで表現する。このグラフは次の2つの要素から構成する。

ノード：URL の集合

アーク：任意のノード間を結んだもの

アクセスモデルは、アクセスログをグラフで表現したものであり、コンテンツモデル同様、以下のような要素からなるグラフとして表現できる。

ノード：ユーザがアクセスした URL

アーク：ユーザがたどったリンク

ユーザモデル構築法としては、コンテンツモデルとアクセスモデルに対し、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。ユーザモデルのグラフは以下の要素からなるグラフとして表現で

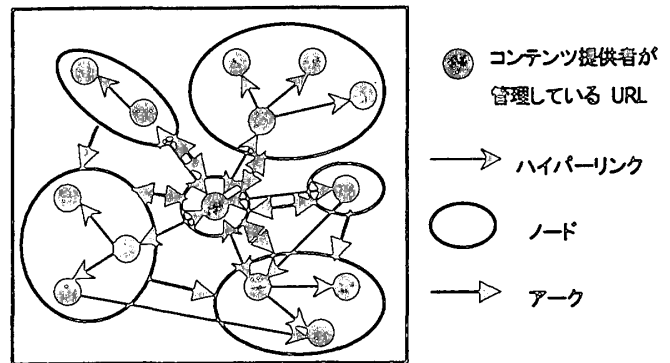


図1 コンテンツモデルの概念図

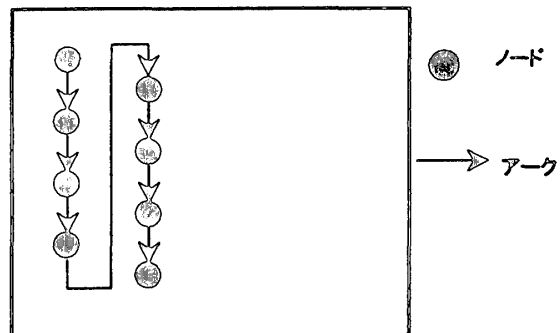


図2 アクセスモデルの概念図

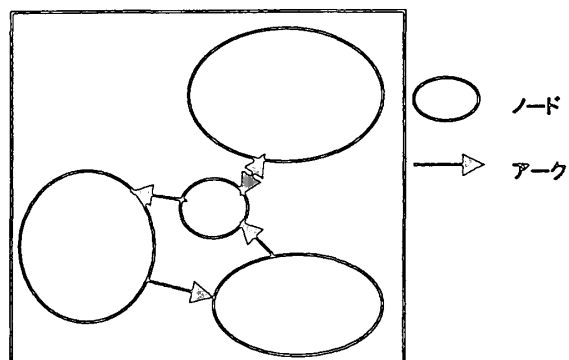


図3 ユーザモデルの概念図

きる。

ノード：コンテンツモデルのノードの全部または一部

アーク：ユーザがノード間をアクセスした場合に、ノード間を結んだもの

まず、コンテンツモデル、アクセスモデル、ユーザモデルのグラフを次のように定義する。

コンテンツモデル

URL 集合 $I = \{v_1, v_2, \dots, v_y\}$

ここで, I は, コンテンツ提供者が管理している URL 集合

コンテンツモデル $C = \langle V^C, A^C \rangle$

C のノード集合 $V^C = \{V_1, V_2, \dots, V_n\}$

コンテンツ提供者の判断により, URL 集合 I を n 個の部分集合 V_1, V_2, \dots, V_n に直和分割する.

ノード $V_i = \{u'_1, \dots, u'_k; u'_d R u'_e, u'_i \in I, 1 \leq d, e \leq k_i\}$
($i: 1 \leq i \leq n$)

ここで, R は, $I \times I$ の部分集合であり, I 上の関係は, “コンテンツ提供者が判断する意味的な関係” である.

C のアーク集合 $A^C = \{a_1, a_2, \dots, a_n\}$

アーク $a_k = (V_i, V_j) \in A^C$

ただし, a_k は, V_i 内の URL と V_j 内の URL とにハイパーリンクが張られているか, V_i と V_j との間に意味的な関係がある場合である.

アクセスモデル

アクセスモデル $L = \langle V^L, A^L \rangle$

L のノード集合 $V^L = \{u'_1, u'_2, \dots, u'_i\} (u'_i \in I)$

ここで, u'_i は, 時系列順に並んだ URL である.

L のアーク集合 $A^L = \{a_1, a_2, \dots, a_{k+1}\}$

アーク $a_i = (u'_i, u'_{i+1}) \in A^L$

ノード u'_i から u'_{i+1} までの時間間隔 t_i をアーク a_i の属性値とする.

ユーザモデル

ユーザモデル $U = \langle V^U, A^U \rangle$

U のノード集合 $V^U = \{V_1, V_2, \dots, V_x\} (x \leq n)$

※ n はコンテンツモデルの位数

U のアーク集合 $A^U = \{a_1, a_2, \dots, a_x\}$

アーク $a_i = (V_i, V_j) \in A^U$

興味・関心度 v_{fi} をノード V_i の属性値とする.

関連度 a_{fi} をアーク a_i の属性値とする.

これに従って, 次のようにユーザモデルを構築する.

(1) 興味・関心度の計算

$$v_{fi} = \frac{\sum_{p=1}^i \delta(p)}{\left(\sum_{q=1}^n \delta(q) \right) \left(\sum_{p=1}^i \delta(p) \right)}, \quad v_{fi} (i=1, 2, \dots, n)$$

$$u'_p (p=1, 2, \dots, t)$$

ここで, $\delta(p) = \begin{cases} 1: (u'_p \in V_i) (\forall u'_p \in V^L, \forall V_i \in V^C) \text{ のとき} \\ 0: \text{その他} \end{cases}$

$\delta(q) = \begin{cases} 1: (u'_p \in V_i) (\forall u'_p \in V^L, \forall V_i \in V^C) \text{ のとき} \\ 0: \text{その他} \end{cases}$

※ n は コンテンツモデルの位数

※ t は アクセスモデルの位数

(2) 関連度の計算

$$a_{fi} = \sum_{p=1}^{t-1} \delta(p) \cdot \frac{u'_p (p=1, 2, \dots, t)}{a_{fi} (i=1, 2, \dots, n)}$$

ここで, $\delta(p) = \begin{cases} \frac{1}{t_p}: (u'_p \in V_i) \wedge (u'_{p+1} \in V_j) \\ (\forall u'_p, u'_{p+1} \in V^L, \forall V_i, V_j \in V^C) \text{ のとき} \\ 0: \text{その他} \end{cases}$

※ t は アクセスモデルの位数

※ n は コンテンツモデルの位数

(3) ユーザモデルのグラフの構築

$V_i \in V^U: v_{fi} > 0$ のときユーザモデルのノードとする

$a_i \in A^U: u_i (v_{fi} > 0) \cap u_j (v_{fj} > 0) (\forall u_i, u_j \in V^U)$ のとき

ユーザモデルのアークとする

4. ユーザモデルの評価実験

調和系工学分野 WWW サイト [2] を対象に, 2001 年 10 月 1 日から 2001 年 10 月 7 日までにアクセスしたユーザの中から 3 人を無作為に抽出し被験者とした. その結果, コンテンツ提供者と個々のユーザとの間に存在するギャップをモデリングできた. また, 他ユーザに共通して存在する特性もモデリングすることができた. 今後, 統計的な処理を可能とするために, より多くのデータを用いて実験を行う予定である.

5. 今後の展開

北海道観光情報 WWW は, 北海道の観光スポット, 宿泊施設, 地図, 観光モデルコース, クチコミ情報などを統合的に提供する. この場合, 構築したユーザモデルを用いて, 例えば, ユーザ特有の関連カテゴリを表示したり, ユーザの行動パターンや興味・関心を考慮した宿泊施設や観光スポットを表示したりする等の個人適応の他に, さらに, 統計的な処理によってユーザの傾向を把握することで, 「コンテンツ自体の改善」などにも利用できる. これによって, コンテンツ提供者は, 自身が提供する WWW についての客観的な評価を得ることができる.

6. おわりに

今後の課題として, 観光情報特有のユーザ特性や, 実際の北海道観光情報ポータルサイトのアクセスログをいかに入手するか等についての検討が必要である.

参考文献

- [1] 三浦, 高橋, 島: 個人適応型 WWW におけるユーザモデル構築法, 情報処理学会論文誌, Vol.39, No.5, pp1523-1535, 1998.
- [2] <http://ses3.complex.eng.hokudai.ac.jp/>