

携帯電話向けサイト構築方法の提案

小林 宏至 田口 雅人 藤岡 秀樹

日立ソフトウェアエンジニアリング(株)

1. 序論

インターネットに接続可能な携帯電話の普及に伴い、携帯電話が企業の業務に活用されつつある。企業は携帯電話を持つ従業員に対し、ポータルサイトを通して様々なサービス(ニュースや在庫確認などの情報提供、旅費清算、勤休管理など)を提供している(B to E)。このような企業のポータルサイトでは、従業員の役職や業務に合わせて、提供しているサービスに複雑なアクセス権が設定されている。

携帯電話は、画面が小さく、入力手段も制限されていることから、コンテンツの表示能力及び操作性がデスクトップ PC よりも劣る。そのため携帯電話向けサイトは、従業員が容易に要求するサービスにアクセスできるように高度にパーソナライズされている必要がある。

本論文では、アクセス制限がある既存のサイトからユーザ(従業員)に合わせた携帯電話向けサイトを構築する方法を提案する。

2. パーソナライゼーション

既存サイトからパーソナライズされた携帯電話向けサイトを構築する方法について説明する。ここでいう既存サイトとは、アクセス制限のあるサイトを指し、携帯電話向けに作成されたサイトでも PC 向けに作成されたサイトでもよい。結果的に得られる携帯電話向けサイトは、ユーザにとって必要なリンクが適切な数だけ表示されているメニュー形式のページから構成される。

2.1 サイトの定義

サイト W を以下のように定義する。 W はページ p の集合 P とリンク l の集合 L から構成される。ページ p は、他のページへのリンクを目的としたリンクページ p_L (ホームページ p_H を含む) と、サービスを提供するデータページ p_D の 2 種類がある。リンク l は組 (p_s, p_d, t, w) として表現する。ここで p_s はリンク元のページ、 p_d はリンク先のページ、 t はハイパーテキスト、 w は重みを意味する。重み w は次節で詳しく述べる。

$$W := P \cup L$$

$$P := \{p \mid p \in W\}$$

$$L := \{(p_s, p_d, t, w) \mid p_s, p_d \in P\}$$

特にリンク元あるいはリンク先のページが p であるリンクの集合をそれぞれ $\overline{LP}(p)$ 、 $\overline{LP}(p)$ とする。また両方あわせたものを $LP(p)$ とする。

$$LP(p) := \overline{LP}(p) \cup \overline{LP}(p)$$

$$\overline{LP}(p) := \{(p, p_d, t, w) \mid p_d \in P\}$$

$$\overline{LP}(p) := \{(p_s, p, t, w) \mid p_s \in P\}$$

2.2 サイト構築手順

特定のユーザ(以下、UserA とする)に対して、既存のサイト W_o からパーソナライズされた携帯電話向けサイト W_p を構築する手順を説明する。ここで述べる方法は厳密解を与えるものではなく、近似解を与えるものである。

Step1 アクセス権の設定

最初に W_o のページを、 p_L と p_D とに分類し、UserA が利用できるサービスを提供する p_D に対して UserA にアクセス権を設定する。UserA にアクセス権がある p_D を \overline{p}_D とする。

Step2 不要なリンクの削除

W_o には UserA にアクセス権がない p_D へのリンクが含まれているため、UserA が必要なサービスへアクセスするのを困難にしている。そこで \overline{p}_D に至らない不要なリンクを削除する必要がある。このような不要なリンクを発見するために、 p_H からすべての \overline{p}_D への最短経路を計算する。どの最短経路上にもないリンクが不要なリンクに相当する。

p_H から \overline{p}_D への最短経路の計算には、リンクの重み w を用いて、Dijkstra のアルゴリズムを使用した。ここで w はそのリンクが最短経路への含まれ易さを意味し、デフォルト値は 1 で値が小さいほど含まれ易くなる。重み w は、例えば以下のような場合に設定する。

(1) FRAME

W_o が PC 向けのサイトの場合、通常 FRAME のページが含まれている。しかし携帯電話のブラウザは FRAME に対応していない。そこで携帯電話向けのコンテンツに変換する場合 FRAME のページを、このページ内に表示する複数のページの一つで代表させる場合がある。このとき、FRAME のページからその代表させるページへのリンクの w を 0 にし、それ以外の FRAME ページからのリンクの w を無限大にする。

(2) 過剰なショートカット

W_o が PC 向けのサイトの場合、一ページに多くのショートカットを含んでいる。しかし携帯電話向けサイトの場合、これら全てを表示するのは過剰である。そこでこれらのショートカットの w を無限大にして、ショートカットが最短経路に含まれないようにする。

以上から得られた最短経路に含まれるリンクの集合を $L_{\min}(\bar{p}_D)$ 、ページの集合を $P_{\min}(\bar{p}_D)$ とする。これをすべての \bar{p}_D について計算し、すべての \bar{p}_D に対する $L_{\min}(\bar{p}_D)$ 、 $P_{\min}(\bar{p}_D)$ の和を、それぞれ携帯電話向けサイト W_p のリンク集合 L_p 、ページ集合 P_p とする。

$$L_p = \bigcup_{\bar{p}_D \in W_o} L_{\min}(\bar{p}_D), \quad P_p = \bigcup_{\bar{p}_D \in W_o} P_{\min}(\bar{p}_D)$$

Step3 ページの統合

Step 2 で得られた W_p は、不要なリンクが削除されているために、一ページに表示するリンク数が調節されておらず、リンク数が1個か2個しかない無駄なページが存在する。このようなページは、結果としてユーザの操作回数を増加させる。そこで類似度の高いページ同士を統合し、一ページに表示するリンク数を調整する。以下、ページ間の類似度の判定およびページ統合方法について説明する。ただし、これらの操作は p_L に対してのみ行う。

(1) 類似度判定

2つのページ p_a, p_b に対する内容の類似度を、そのページへのリンクとそのページからのリンクから判断する。ページ p_b のリンク集合 $LP(p_b)$ において、要素 p_b をもう一方のページ p_a に置き換えたリンク集合を $LP(p_b \mapsto p_a)$ とする。ただし、以降リンクの重み w は使用しないのでリンクの組から削除する。

$$\begin{aligned} LP(p_b \mapsto p_a) &:= \overline{LP(p_b \mapsto p_a)} \cup \overline{LP(p_b \mapsto p_a)} \\ \overline{LP(p_b \mapsto p_a)} &:= \{(p_a, p_d, t) \mid p_d \in \overline{LP(p_b)}\} \\ \overline{LP(p_b \mapsto p_a)} &:= \{(p_s, p_a, t) \mid p_s \in \overline{LP(p_b)}\} \end{aligned}$$

このとき p_a, p_b 間の類似度を表す距離 d を、 $LP(p_a)$ と $LP(p_b)$ のリンク数の少ない方を、 $LP(p_a)$ と $LP(p_b \mapsto p_a)$ の共通のリンク数で割ったものとして定義する。ここでページ p の総リンク数は $LP(p)$ の組数 $|LP(p)|$ で表される。この距離をすべてのページ間について計算する。

$$d := \frac{\min(|LP(p_a)|, |LP(p_b)|)}{|LP(p_a) \cap LP(p_b \mapsto p_a)|}$$

(2) ページ統合

(1) で求めた距離 d の最小となるページ同士を統合する。統合したページを p_c としたとき、 W に対して以下の操作を行う。

$$\begin{aligned} P &= P - \{p_a, p_b\} \cup \{p_c\} \\ L &= L - LP(p_a) - LP(p_b) \\ &\quad \cup LP(p_a \mapsto p_c) \cup LP(p_b \mapsto p_c) \end{aligned}$$

(1)と(2)の操作をすべての $p \in W_p$ に対しその p に表示するリンク数 $|\overline{LP(p)}|$ が設定された範囲に収まるまで繰返す。その後 p を集合 $\overline{LP(p)}$ のすべての要素を表示するメニュー形式の携帯電話向けページとして定義しなおす。

3. 関連研究

アクセスしてきたユーザや機器の属性に合わせてコンテンツを変化させる従来の技術(コンテンツアダプテーション)には主に2種類存在する。

(1) ページに対するアダプテーション

ページのインターフェイスを変更する技術である。この技術には、内容とプレゼンテーションとロジックとを分けてページをXMLで記述することによりアダプテーションを行うUIML[1]等や、既存コンテンツの構成要素に対して重要度や役割をRDFで記述することによりアダプテーションを行うAnnotation[2]がある。

(2) サイトに対するアダプテーション

サイトのリンク構造を変更しユーザを特定のコンテンツにナビゲートする技術である。この技術には、特定ユーザのアクセスログを解析し、ユーザが頻繁に利用するページへナビゲートするためにリンクの追加や削除を行うPROTEUS[3]や、アクセスログからユーザをクラスタリングして、リクエストの多いページへのショートカットを追加するMINPATH[3]がある。

本研究の方法は、従来のサイトに対するアダプテーションでは扱えないアクセス権が設定されたサイトを対象とし、サイト全体から不要なリンクを削除し、ページ内のリンクを決めている点で、ページ単位の制御しかできないページに対するアダプテーションよりも、パーソナライゼーションに関してより大きな効果が期待できる。

4. 結論

本論文では、アクセス権が設定された既存サイトからパーソナライズされた携帯電話向けサイトを構築する方法を提案した。本方法は、ホームページからアクセス権のあるページまでの最短経路を求めて不要リンクを削除し、さらに類似度の高いページ同士を統合しページ内のリンク数を適切なものにしたことにより、ユーザの容易なサービスへのアクセスを可能にした。

参考文献

- [1] Marc Abrams et.al : "UIML: An Appliance-Independent XML User Interface Language", Proceedings of the 8th World Wide Web Conference (1999)
- [2] Masahiro Hori et.al : Annotation of Web Content for Transcoding : W3C Note (July 1999) : URL= <http://www.w3.org/TR/annot/>
- [3] Corin R. Anderson : A Machine Learning Approach to Web Personalization. Ph.D. thesis. University of Washington, Department of Computer Science and Engineering. (2002)