WAP Management with user defined content reduction
Prof. D. B. Kulkarni, Amar R. Buchade (M.E. CSE)
Department of Computer Science and Engineering,
Walchand College of Engineering, Sangli, India.
d_b_kulkarni@yahoo.com, buchadear@rediffmail.com

Abstract
In today’s world internet is the main source of information. WAP users can not read most contents designed for PC users via their mobile phone screens. For those sites that have been maintained with HTML, time and manpower costs will be incurred to rebuild them with WML. The WAP site management system is proposed for this purpose. The web sites can be automatically translated to proper WAP contents based on parameters like keywords, compactness factor and depth of content search provided by mobile user. Management system show relevant information and retains those links that having content related to keywords and reorganizes it to fit size of mobile client screen.

Keywords: WAP, WML, WBMP, content reduction, link retaining.

1. Introduction:
Conventionally, the vast majority of users access the Internet using desktop computers. The WAP forum, a consortium founded by a group of cellular phone manufacturers and software companies such as Ericsson, Motorola, Nokia and Phone.com established a stack of wireless application protocols in 1997 to facilitate Internet services to digital cellular phones and other wireless handheld devices.

Internet content providers can provide information and services to mobile phone users in WML descriptions. For mobile users, wireless communication is capable of providing instant information and services such as E-mail service, real time stock information, banking news etc.

Ordinary web technology can not be used here because handheld devices have less memory, limited screen size and for most Internet Service Providers (ISP), the task of reorganization of web content for wireless appliances of user is highly time consuming. Maintenance of an additional WAP site is also burden to Internet content providers. To tackle these problems, WAP site management system can be implemented by ISP’s. WAP site management translates HTML documents into WML documents, images into proper type and reorganizes the contents. WAP system also reduces volume of contents by employing a content reduction algorithm and retains those links that having relevant information.

2. Relevant Research:
The workgroup WEB2WAP has been founded in February 2002 at the University Of Technology & Economics in Budapest. This group has implemented HTML-to-WML system, but it could not handle tags properly [7].

The limitation such as communication bandwidth, memory size, and screen size of mobile device puts constraints on accessing and displaying HTML contents without any alterations. The WAP site management system implemented in this paper is capable of translating HTML documents into a feasible WML format by reorganizing and displaying the information satisfying the user defined criterion.

3. WAP site management system
The WAP site management structure shown in Fig. 1 consists of following modules.

a. Web Content Analysis: It checks the web contents for type (HTML/WML). WML contents are sent to the mobile device without translation and the HTML contents are sent to content reduction and link retaining module.

![Figure 1: Structure of WAP system](image-url)
b. Content Reduction and Link Retaining Module: This module extracts relevant information from the web page with the help of user specified keyword and other parameters with retaining the relevant links.

c. Translation Module: This module translates reduced content in HTML format to WML format and images within into WBMP format. It consists following modules.

1. HTML to WML-This module considers WML rules for translating HTML-to-WML. It eliminates HTML tags that are not supported by WML e.g. tags of script, style sheets, font and comments. It also handles tags like A, FORM, FRAMESET, TABLE and converts them into suitable form [2][3][4].

Links- Links are placed in tag A. Suppose URL http://www.mysite.com/index.html contains following links.

- `<a href="http://www.site.com/page1.htm">Hi</a>`
- `<a href="http://www.site.com/page1.htm">Hi</a>`
- `<a href="http://www.site.com/page1.htm">Hi</a>`
- `<a href="http://www.site.com/page1.htm">Hi</a>`

It is processed as:

- `<a href="http://www.site.com/page1.htm">Hi</a>`

User input- Mobile users should be able to submit forms to the server. Some commonly used elements such as checkbox, text area are not supported by WML. The conversion of user input is shown in the following table.

<table>
<thead>
<tr>
<th>HTML Text</th>
<th>password</th>
<th>radio</th>
<th>checkbox</th>
<th>reset</th>
<th>Submit</th>
</tr>
</thead>
<tbody>
<tr>
<td>WML Input</td>
<td>Input</td>
<td>select</td>
<td>select</td>
<td>anchor</td>
<td>anchor</td>
</tr>
</tbody>
</table>

Frameset- Mobile devices do not provide support for showing several documents simultaneously in separate parts of a screen. Thus menu with a link to each of the several documents in the frameset is created.

e.g.

```html
<frameset rows="20%, 60%, 20 %">
    <frame src="http://www.t.com/p1.htm" name="one">
    <frame src="http://www.t.com/p2.htm" name="two">
    <frame src="http://www.t.com/p3.htm" name="three">
</frameset>
```

WAP site system translates it into following code:

- `<a href="http://www.t.com/p1.htm">one</a>`
- `<a href="http://www.t.com/p2.htm">two</a>`
- `<a href="http://www.t.com/p3.htm">three</a>`

**Table** - For converting HTML Table statements, following rules are taken into consideration.

1. If number of columns is less than two; HTML Table directly translated into WML Table.
2. If numbers of columns are greater than two, HTML table translated into set of options.

Image – All pictures have to be converted to a format suitable for being displayed on cell phone displays. Most of cell phone support WBMP format. Any color in the JPG, GIF or PNG image is converted to black and white.

Figure 2: Image Converting Process

Figure 2 shows how pictures are converted to WBMP format. The image is first loaded to Java Image class and scaled to desired format. The user can set scale factor manually or one of the predefined scales can be used. A process called “dithering” is used to translate colors into different densities of black dots on a white background, thus producing different shades of gray. [5][12]

2. Content Reduction and Link Retaining Module:

This module extracts relevant information from the web page and retains the links that contains information related to user specified keywords.[9][10][11]

3. Content Reorganization:

The content reorganization module analyzes the structure of translated WAP contents to reorganize it for display on mobile client.[6]

4. Content Reduction and Link Retaining Algorithm: A content reduction algorithm proposed in [9] is used to reduce contents of the documents originally designed for personal computers. The user can specify keywords and the sentences that are irrelevant to the keywords will be removed from the document. For the web page links that contain keywords should also be
The proposed algorithms allows user to specify the depth, which helps in retaining/rejecting the link based on presence/absence of keyword in the link-content. The user has given request URL, keyword and depth. After downloading the page, if it contains specified keyword, link of the page is preserved. Similarly, if the page does not contain link that contains the keyword, it is removed. Thus algorithm recursively follows links on the page to find keyword up to specified depth.

5. Experimental Results and Evaluation:

WAP management system is developed in JAVA on Windows 2000 platform and Nokia Internet Toolkit 3.0 used for demonstration. In content reduction and link based algorithm, user can specify parameters such as keywords, compactness factor to obtain the relevant information. The results of the system developed are shown with the figures and the interpretations as follows.

Figure 4, Figure 5 shows the experimental results for political, economic, sport and technical documents based on various factors.

5.1. Effects of keywords on documents:

Figure 4 indicates that as the document contains more keywords, percentage of removal of words decreases. Thus performance of the content reduction algorithm is mainly affected by keyword population of the documents.

5.2. Effects of Compactness on Documents:

It can be observed that the number of sentences in the reduced document decreases as Compactness factor increased. The document processed will be still readable because the sentences that contained keywords would have been preserved.

5.3. Response time of System:

Figure 6 presents the experimental results on document based on no. of keywords.

It can be seen that response time increases as no. of keywords entered by user.
increases because of time required to search each keyword in sentence.

Figure 7 presents the experimental results on document based on compactness factor.

![Figure 7: Response time based on compactness](image)

It can be seen that response time varies as Compactness factor increases. This depends on the parameters specified by the user.

6. Conclusions:

I-WAP system designed automatically converts HTML contents to WML format, converts images into WBMP format, reduces the original contents of a Web page with user specified parameters and retains the links that contains specified keywords at the specified depth of document hierarchy.

7. Future Work:

Current work can be extended as auto-computation of factors such as compactness; keywords will give no. of pages/no. of words/no. of lines depending on the user requirement. I-WAP System can be made to personalize the response to specific mobile client based on the capabilities of mobile device with pre-registration of model by mobile user with ISP. Implementation of cache management will improve response time of system.

8. References:


3. Scott Tilley, Bas Toeter, Ken Wong ,”Issues in Accessing Web Sites from Mobile Devices”


6. Staffan Björk, Telia Mobile Rolf Danielsson Telia Research Swedish Institute for Computer Science, ”WEST: A Web Browser for Small Terminals”

7.http://www.mcl.hu/web2wap/ Web2WAP workgroup site:


Acronyms:

- **PC** - Personal Computer.
- **HTML** - Hypertext Markup Language.
- **WML** - Wireless Markup Language.
- **WBMP** - Wireless Bitmap.
- **ISP** - Internet Service Provider.