

Client Device Based Content Adaptation Using Rule Base

Velammal and Anandha Kumar
Department of Computer Science and Engineering,
Anna University, Chennai, India

Abstract: Problem statement: Content adaptation have been playing an important role in mobile devices, wherein the content display differs from desktop computers in many aspects, such as display screens, processing power, network connection bandwidth. In order to display web contents appropriately on mobile devices and on other types of devices such as hand computers, PDAs, Smart phones, it is important to adapt or transcode them to fit the characteristics of these devices. **Approach:** Existing content adaptation systems deploy various techniques which have been developed for specific purposes and goals. By exploiting various possible combinations of available resources, appropriate adaptation process can be carried over on the actual data, so that the information can be assimilated in a different end system other than the intended system. In this study, we present a content adaptation system based on rules created for mobile devices. Rules are invoked based on the individual client device information. **Results:** The adaptation has been performed according to the delivery device which was formalized through the profiler system. A profile holds information about the hardware and software specifications of the device thereby enabling the adaption of web content based on their characteristics which enables the user to access the web easily on various devices. **Conclusion/Recommendation:** This study enhances the viability of the information being presented to user, which will be independent of the end system being used for accessing the information. With the help of configurable rules, effective content adaptation can be achieved to provide optimal result.

Key words: Content adaptation, mobile device, rule base, page formatting, profiler system, mobile devices, existing content, adaptation systems, client device

INTRODUCTION

Content adaptation is the action of transforming content to adapt to various device interfaces. This procedure is usually related to mobile devices that require special handling because of their limited computational power, small screen size and constrained keyboard functionality.

Advances in the capabilities of small, hand held devices such as mobile phones (cell phones) and Personal Digital Assistants have led to an explosion in the number of types of devices that can now access the Web. Content adaptation is one approach that can solve this problem. Rather than requiring authors to create pages explicitly for each type of device that might request them, content adaptation transforms an author's materials automatically.

In the dynamic content Adaptation to the mobile device, Web users face challenges when trying to enable users to browse the World-Wide Web from handheld devices such as Mobile, Palm Pilots and various other communication devices. Among these challenges, screen size limitations and adaptation of content for structure and pointer Object and other

information stand out. Screen size limitations are an issue because most HTML pages are designed to be viewed on desktop displays. Their page layout assumes that users can see large portions of each page at once. The much smaller page excerpts displayed on any one handheld device screen can interfere with users' comprehension and the scrolling activity is always a time consuming. This results overhead in obtaining the required content.

Related works: While accessing the web resource through the internet, Integration of content and context is important because it is crucial in human-human and the human-computer communication. It is difficult for humans to recognize various objects in web page of different sizes. There have been issues related to bridging semantic and the conceptual gaps in the past few years. Our intention with this special issue is to go beyond the specific research topics in the field and to focus on an over-arching issue and that is, how best desktop content is adapted to the mobile device with an effective content exploit context in content analysis.

Corresponding Author: Velammal, Department of Computer Science and Engineering, Anna University, Chennai, India

In the last few years, tons of research study in content adaptation for content-type-specific and general-purpose content adaptation systems have been developed. Distributed architecture supports metadata extraction by exploring interaction mechanisms among users and content. The interaction activity addressed in this study is related to peer-level annotation, where any user acts as an author, thereby being able to enrich the content by making annotations (Manzato, 2009). In addition to enrichment of content, research has also been performed in monitoring the performance characteristics such as response time, percentage of successful completion, timeouts and external errors with the help of dynamically composed media web services (Iskandarani, 2010).

Only the services preferred by the user have been made portable with the help of session initiation protocol across various networks for achieving seamless service mobility. A collaborative caching system in cluster environment (Elfaki *et al.*, 2011) has been studied to increase the performance of data retrieval by introducing caching mechanism. The local cache hit ratio and the delay has been investigated to improve the responsiveness and reduce the number of hops between requested node and source node.

Advancements in technology lead to enhance the learning process at all possible levels. The user can use mobile technology for m-learning (Nasiri and Deng, 2009) which permits the users to access learning materials through any mobile devices like WAP enabled mobile phones, Palmtops and other mobile peripheral devices. With the help of Mobility Prediction (Ghosh and Bhaumik, 2010) movement of the user can be predicted which will help in providing location based services to the mobile users and in this case, predicting next location of the user will absorb considerable amount of time in responding to the actual request by the client.

Web pages often contain clutter (such as pop-up ads, unnecessary images and extraneous links) around the body of an article that distracts a user from actual content. Hence it is necessary to remove clutter thereby making content more readable by changing font size or removing HTML and data components such as unnecessary images, which eliminates the webpage's inherent look and feel. "Content reformatting" aims to reproduce the entire webpage in a more convenient form. The solution directly addresses "Content Extraction". One method is to study with the Document Object Model tree, rather

than with raw HTML markup. Implementation of caching mechanism to access the data is an efficient manner. When applying the Selective Adaptive Sorted (SAS) cache invalidation strategy (Safa *et al.*, 2010) minimum level of false invalidation strategies obtained. The performance matrices evaluated and compared with the selective cache invalidation strategy and the updated invalidation report strategy. To display these contents appropriately on mobile devices, content must be adapted or transcoded to fit the characteristics of these device. The adaptation is performed according to the delivery context information (Luo *et al.*, 2009) which has been formalized by means of a profiler system. A profile holds information about the specific access device, the user preferences and the device working condition. This will provide some automatic means to convert any existing content into a version suitable for rendering in a client device requesting for that content.

A DOM-based content extraction approach is used to extract useful content, relevant contents and remove distracting features from HTML pages. In this approach, images and links are the targets for removal. The system can efficiently remove ads, banners and other noisy contents. However, in some cases, images and links contain useful and important information. Removing them without further investigating their roles can have a negative impact on the browsing experience. Also, this system does not fully provide general-purpose content adaptation techniques. In this approach, a decision engine has been designed to find the optimal adaptation based on QoS attributes. They try to quantitatively measure the QoS and some negotiation algorithms are used to determine the optimal content version. The adaptations in these studies focus on basic objects, such as text and images (Ghosh and Bhaumik, 2010).

With Frame resizing algorithm using DCT (Velammal and Kumar, 2010), images can be adapted suitable to the requirement of the devices. The caching of most recently accessed images with the help of Maximum profit replacement algorithm, enhances the retrieval of contents much faster. However this has been applied only for image adaptation in mobile devices and has to be extended for other temporal characteristics of multimedia content prevalent in the web.

MATERIALS AND METHODS

System architecture: The system developed should be able to adapt the content. The below architecture has been developed as a java web application.

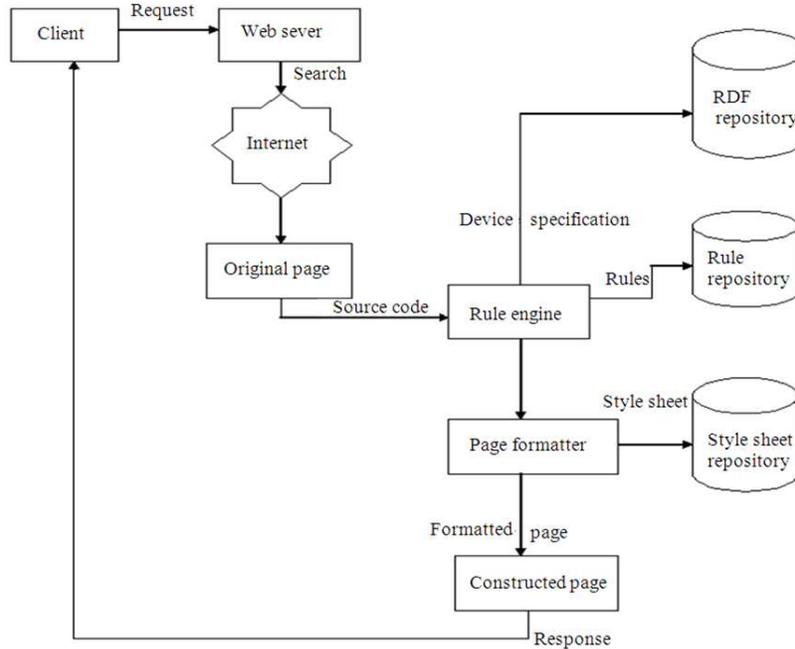


Fig.1: System architecture

```

<rule name="rule1" description="if: X
Is human then: X is mortal" >
  <if left Term": X" op="=" right
Term="MSI" />
  <then method="prop. set clause"
Arg1="style.css"/>
</rule>
  
```

Fig. 2: General rule in rules repository

The block diagram explains the various steps involved in the content adaptation, which include RDF Profiles repository, Rule Repository, Rule Engine, Style sheet repository, Page Formatter and web server to retrieve the requested content from web. The various modules involved in the system and their interaction has been depicted in Fig 1.

The request from the client is sent to the web server and it performs a search to retrieve the information from the internet. The required web page is then sent to the rule engine which finds information applied to the original web page so it can be viewed perfectly using any kind of device.

Usually when web content designed to be displayed on desktops are adapted for the small screen size device, the content often looks messy making it

difficult for users to read the contents, hence unwanted information have to be discarded and also certain things like image formats need to be converted since one device may support one format and other may support another. Hence adaptation need to be carried over.

Creating repositories: This module focuses on creating repositories. The Rule Repository frames rules for the content analysis and adaptation. The content adaptation rules define how to adapt various type of content depending on the device types as shown in Fig. 2.

User database contains the information about the user which will help the registered user to access the website. User database contains the user's general personal information and the details about the type of the device frequently accessed by the user. The device specification is extracted from a RDF repository which is performed with the help of jena. Jena is an API which allows one to parse, create and search RDF models. Jena takes an XHTML page that contains embedded RDF, extracts and parses the RDF. This is done with the read() method in the Model interface. The extracted data is stored in a table which is used to check the software and hardware specification of that particular device.

Page formatter: This module deals with managing results and sending them to a page formatter where it

sets some basic values like dimensions of a new image or its values of scale for a runtime adaptation of content to be presented. Thus this will format the page in accordance with the user device. In this study, jsoup is used for formatting the requested user web page based on the device screen size. The screen size of the device is extracted from RDF repository. Jsoup is a Java library for working with real-world HTML. It can parse HTML from a URL, file, or string. It can find and extract data, using DOM traversal or CSS selectors. The HTML elements, attributes and text can be manipulated. It can clean user-submitted content against a safe whitelist. jsoup is designed to deal with all varieties of HTML found in the wild, from pristine and validating to invalid tag-soup; jsoup can create a sensible parse tree. With the help of this, pictures or images can be formatted easily based on the device screen size.

Algorithm:

Step 0: Get the page content:

Get requested page and retrieve the page source code of the requested page

Step 1: Extract the multimedia content:

Check whether it has text images or other multimedia content.

Step 2: If it has images extract its height and width.

Step 3: Now check for device type if it is small screen device then reduce the height and width of the image accordingly.

Step 4: If it is not small screen device then don't change the height and width of the image.

Rule engine: Feature vectors were extracted from the filtered image and fed as inputs to the first layer of neurons of the multilayered neural network. The result obtained from the output layer will be compared to the target output. Based on this, the network adjusts the weights for each neuron at the output and hidden layer respectively (back-propagation learning method). The network then proceeds to the next image, compares the result with the targeted output and adjusts the weights again until a pre-set goal is met.

In this study java rule engine JRule Engine is created. The rule engine is a if/then statement interpreter. If/then statements that are interpreted are called rules. The input of this rule engine will be rule execution set and some data output will be selection of rules that needs to be implemented.

Algorithm:

Step 0: Get the rule service provider from the provider manager and get the rule administrator.

- Step 1: Get an input stream to a test XML rule set and parse the rule set from the XML document
- Step 2: Register the rule execution set and ret a rule runtime and invoke the rule engine
- Step 3: Create a stateful rule session and the rule to it.
- Step 4: Then Fire all rule that is execute the rules
- Step 5: If matches with the rule in rule repository then accordingly apply the style sheet

RESULTS

Ultimate goal of the proposed system is to provide the capability to adapt this content to meet the specific needs of any number of user groups. This enables the user to adapt the web content based on the type of user device. This will dynamically identify the type of device from where the request is coming and accordingly it will adapt the web page so that the user is able to access the Internet. Thus it provides user friendly content to the mobile device users.

The expected adapted web page given below (Fig. 3) represents the adapted page without removing the unwanted content in the desktop. The Fig. 4 represents the adapted page without removing the unwanted content in the mobile which leads to a messy look. The user finds difficulty in reading this. The Fig. 5 represents the final adapted page in desktop. The Fig. 6 represents the final adapted page in the mobile with necessary contents.



Fig. 3: Formatted page with unwanted content



Fig. 4: Formatted page with unwanted content



Fig. 5: Expected web page in desktop



Fig. 6: Expected page in mobile

DISCUSSION

A rule engine in tandem with rule base has been used in the system to dynamically carry over the content adaptation policy based on the device. By having a general base for the characteristics of all the communication devices in the rule base, the effectiveness of the adaptation has been increased manifold due to the customized policies. The rule base can also be equipped to hold user based adaptation policies based on their own preferences. This greatly helps the user in obtaining the required content in their most desirable area of interest and in the preferable format. The style sheet repository helps to reproduce the requested content in the most suitable presentation format according to the communication device. The user experience is greatly enhanced by implementing this presentation layer adaptation over the requested content.

CONCLUSION

In this study, we have designed a Dynamic content adaptation policy which process the input HTML file to extract the page source and adapt based on their content type. This system is able to find the user agent that is user

device from which the page is requested, so that the web page is dynamically adapted based on the user device. Hence we can dynamically get the user requested page based on user device.

This study was developed to adapt the web page content dynamically based on the user device. Some of the scope to enhance this system further are as follows.

Improving system efficiency: The current techniques can dynamically adapt the web page content based on the user device; it can be performed only at an average speed. Implementation in a real-time environment requires the system to be hard real time. This can be achieved by implementing a system with higher configuration in terms of memory, CPU speeds.

Making the system more users friendly: The current system adapts the web page content. The system can be made more user friendly by being interactive with the user i.e., when a system recognizes a user device, it can provide an option to the user if he wants to include a new device so that the content adaptation can be done dynamically. If a system assesses that a user device is new, then it will be able to retrieve and include the necessary information so that it can perform content adaptation for the new device.

REFERENCES

- Elfaki, M.A., H. Ibrahim, A. Mamat and M. Othman, 2011. Collaborative caching architecture for continuous query in mobile database. *Am. J. Eco. Business Admin.*, 3: 33-39, DOI: 10.3844/ajebasp.2011.33.39
- Safa, H., H. Artail and M. Nahhas, 2010. A cache invalidation strategy for mobile networks. *J. Network Comput. Appl.*, 33: 168-182, DOI: 10.1016/j.jnca.2009.08.003
- Velammal, B.L. and P.A. Kumar, 2010. Dynamic Content Adaptation for Multimedia Content in Handheld Devices. *Multimedia Tools Appl.*, 51: 379-396, DOI: 10.1007/s11042-010-0669-x
- Iskandarani, M.Z., 2010. A novel approach to signal detection of sensor array units using 5-3-1 rule based matched filter algorithm with intelligent identifiers. *Am. J. Eng. Applied Sci.*, 3: 427-432. DOI: 10.3844/ajeassp.2010.427.432
- Ghosh, K. and K. Bhaumik, 2010. Complexity in human perception of brightness: A historical review on the evolution of the philosophy of visual perception. *online J. Biological Sci.*, 10: 17-35. DOI: 10.3844/ojbsci.2010.17.35
- Nasiri, A. and G. Deng, 2009. Environmental factors influence on mobile learning business. *Am. J.*

Applied Sci., 6: 1225-1234, DOI:
10.3844/ajassp.2009.1225.1234

Luo, J., A. Hanjalic, Q. Tian and A. Jaimes, 2009.
Integration of context and content for multimedia
management. *Multimedia IEEE Trans.*, 11: 193-
195, DOI: 10.1109/TMM.2008.2009179

Manzato, M.G., 2009. User-centered multimedia
content personalization: A services composition
based approach. Mathematics and Computing
Institute.