WebPage Transclusion – An Adaptive Algorithm for Web Technologies

Vishwas J Raval, Amit R Thakkar, Amit P Ganatra and Yogeshwar P Kosta, Member, IEEE

Abstract—As every thing in terms of services or an activity, be it of economic interest or otherwise collapses on the www – for want of data or info, filing an income tax return, submitting an application or even working as a part of a virtual team/company, it is become important, keeping in view of the diversity of users – from educated to the uneducated - methods and mechanisms be discovered/evolved to make the www a better and habitat place providing a conducive environ – in plain words: human user friendly and welcoming - while browsing or laboring for the same as our society transcend into the next century of virtual world and reality. The present scenario – during extraction of data or info, in a www environment, when a user wants to include some text or image from another website for reference – Transclusion is used. Transclusion is the inclusion/incorporation of part of a document into another document through instantaneous reference. Consider for example, an article about a country might include a chart or a paragraph describing that country’s agricultural exports from a different referenced article about agriculture (pointer), rather than copying the included data and storing it in two places. A transclusion embodies modular design, by allowing it to be stored only once (and perhaps corrected and updated if the link type supports the same) and viewed in different contexts. The reference also serves to link both articles. However, this has several limitations which have been addressed by our research and we proposed a new technology – WebPage Transclusion” to include web pages that help to eliminate the above limitations.

Index Terms—HTML, DHTML, Frame, Hyperlinks, Form, Transclusion, Browser, Frame-Bust, Frame-Bust Protection, IFrame

I. INTRODUCTION TO TRANSCLUSSION

Nelson was the first to coin the words "hypertext" and "hypermedia" including "transclusion" in his book, Literary Machines, 1982. Part of his proposal was the ideas that micropayments could be automatically extracted from the reader for all the text, no matter how many snippets of content are taken from various places [1]. Figure-1 illustrates Transclusion.
Nelson has delivered a demonstration of Web transclusion, the Little Transquoter (programmed to Nelson’s specification by Andrew Pam in 2004–2005). It creates a new format built on portion addresses from Web pages; when dereferenced, each portion on the resulting page remains click-connected to its original context—always a key aspect of transclusion for Nelson, but missing in most implementations of transclusion.

II. PROBLEMS ASSOCIATED WITH TRANSCLUSION

At present, transclusion in HTML is somewhat limited by lack of standards support in web browsers. Although all graphical browsers can transclude an image, including a document is a bit more difficult. There are currently two methods of achieving this result:

- The IFrame (inline frame) element that includes whole document/s.
- The Object element, in Firefox and other browsers, permits true document transclusion, enabling one page to be built from several smaller documents.

These two methods mentioned are limited to small images or text only and if one wants to include the whole web-page/s as a part of any other web page, then the same can be achieved through hyperlinks in HTML. See figure 3. [2][3][4][5][6][7].

However, in the case of hyperlinks, when one wants to refer to a particular page, user needs to open that page in a separate window/tab. In case the user opens it in the current operating page-window then the operating page-window gets masked by the particular (referred) page-window and thus the user loses the navigation controls embedded/contained in the window/frame during browsing ultimately paralyzing the browsing process, midway resulting in loss of link (break) and information and difficulty to navigate back to the original start point. This scenario has been demonstrated in figure 4.

III. WEBPAGE TRANSCLUSION TECHNOLOGY

WebPage Transclusion is the technology to include whole web pages across www for references or use. In this paper, our team report the experimental development of an adaptive algorithm called: “The WebPage Transclusion Algorithm” that is sufficiently robust and experimentally verified to take care of the problems faced during web browsing discussed.
above. This algorithm works behind the screen like a daemon while the user browses through www pages in the internet.

IV. THE WEBPAGE TRANCLUSION ALGORITHM

The following steps are conceptual to the design of the WebPage Transclusion Algorithm:

1) Create a web page using frame1, here we divide the window into two parts viz: Navigational control and Content frame. Figure-5 shows the left/top of the window to work as Navigation Frame and right/bottom part as Content Frame - when user starts browsing.

2) Scan the web pages that user browses through, opening it into the content frame by setting target of the web page being browsed equals to name of content frame (pointer).

3) If web page contains hyperlinks extract those links and add it into the navigation frame as a subset of link which has been added in step-2.

4) If web page contains <form> tag, then the Call command for Frame-Bust Protective Algorithm is invoked in order to prevent Frame-Busting. [8]

5) Repeat steps 2 to 4 till browser’s instance is destroyed.

6) Exit.

Figure 6 shows the WebPage Transclusion Process.

A. Advantages

1) As the pages would be opened into the content frame, users would never loose navigation.

2) Users would be able to access all the desired websites using a single browser window, he/she would not be required to open a new tab even.

3) In case of Frame-Busing, with The Frame Bust Protection Algorithm user would feel transparency during the web browsing. [8]

B. Disadvantages

1) However, implementation point of view, it would be difficult to traverse all the pages which are received in response from various web servers and retrieve the links and add it to the navigation frame. For this some efficient algorithm would be required.

2) If Frame-Bust Protection is used to copy contents for Bust-Protection would result in violation of copyrights for which prior permission would be required [6][8]

V. FUTURE WORK

The algorithm is still under implementation phase. We are implementing it in JAVA so that there would not be any issue of platform independency & licensing.

ACKNOWLEDGMENT

We are thankful to The Omnipotent GOD for making us
able to do something. We express our gratitude to the management of CHARUSAT; Shri Charotar MotiSattavis Kelavani Mandal, for providing us research opportunities and their wholehearted support for such activities. Finally, our acknowledgement cannot end without thanking to the authors whose research papers helped us in making this research.

REFERENCES


