IJCAT - International Journal of Computing and Technology, Volume 1, Issue 9, October 2014
ISSN : 2348 - 6090
www.IJCAT.org

Web-Page Complexity and Optimization Mechanism to Reduce Web-Page Load Time

[1] Omkar Sawant, [2] Sachin Godse

[1, 2] Department of Computer Engineering, Sinhgad Academy of Engineering, Kondhwa (University of Pune), Pune, Maharashtra, India

Abstract - Webpage load time is a critical issue in today's world of Internet. It directly affects the user's satisfaction. Recent studies have proposed the concept of 'Web Page Complexity'. Complexity of a website eventually affects the web page load time. A set of metrics to characterise the complexity of a Website has already been formulated. The complexity of a website can be measured at content level and service level. Based on this fact, metrics such as 'number of objects requested', 'MIME type of the objects', 'number of servers used' and 'non-origin contribution' have been identified to be important. While there have been extensive study regarding websites and related issues, little work has been done on understanding the complexity of an individual website and how it is measured and manipulated to improve the user's experience and reduce webpage load time. This paper tries to fill this gap by utilising the concept of 'Complexity of website' and the metrics proposed by prior studies to measure and analyse the complexity of a website. For that purpose, an Analyser tool is proposed which will help determine the complexity of a website in terms of the complexity metrics mentioned above. The analyser tool will generate a report containing information regarding the nature of complexity of a website. Recent studies have mentioned the need to explore more systematic tools and mechanisms to improve page load time. Extending this line, this paper also aims to study the impact of optimization mechanisms such as combining external JavaScript, combining external CSS, minimizing redirects and minify-js to reduce the complexity of a website thereby reducing the page load time. This is to be done by balancing the trade-off between page load time and the complexity of a website because measures taken to reduce page load time should not affect the user's experience and web publisher's revenue.

Keywords **– MIME, Non-origin contribution.**

1. Introduction

World Wide Web is the technological trend that has changed the traditional character of every field in the life of mankind and it has been doing so with continuously upgrading itself to adapt to the changing situations and user's requirements. There hardly is any field that has not been touched with the evolution in the web and its direct or indirect implications. During transition of Web 1.0 which characterised simple static text based information to the web 3.0 which represents rich client and dynamic data processing; World Wide Web has significantly changed in its form and essence [1]. With this changing nature of World Wide Web, new demands and requirements are arising from the users and the system itself.

Web site is the entity which the system of World Wide Web stands upon. The nature and orientation of websites has drastically changed during the evolution of the Web. These new changes have timely added new functionality to improve the client performance as well as user's experience. But these changes on the other have made the websites more and more complex and caused other challenges that need to be considered and tackled to move towards the next step of the evolution of the Web.

One of the most important of these challenges is the increasing complex nature of websites which ultimately causes increase in the webpage load time. Webpage load time is the time taken by the browser to display the page requested by the user completely. As a website grows more and more complex its page load time increases simultaneously [2]. This fact eventually deteriorates user's experience and satisfaction, which are the essential parameters that determines the success of any technology.

Webpage Complexity:

A prior study has proposed the concept of webpage complexity [2]. The term refers to the complex nature of the web page in terms of certain factors. The factors considered are as follows.

1. Number and type of objects requested to load a webpage
2. Number of servers contributed to load a webpage

Based on these parameters; complexity of website is determined in terms of certain metrics. These metrics are as follows.

1. Number of objects
2. MIME type of objects
3. Number of servers used to load the page
4. Non-origin contribution to load the webpage

By applying above theory, it is found that the complexity of a website is a dominant factor that ultimately determines the webpage load time [2]. Considering this fact, this paper presents some optimization mechanisms to reduce the webpage load time. This will be a crucial step in the evolution of Web, as the more the websites become complex the longer the time it takes to load it. It is the need of the time to improve the client performance in spite of its complex nature and rich presentation without disturbing user's experience and satisfaction.

The paper also proposes an Analyzer tool to measure the complexity of a website. This tool will be helpful to determine the nature of the content, structure and orientation of a website which are the factors that need to be considered while implementing optimization mechanisms. Optimization is not a straight forward task. Hence it may differ according to the nature and type of the website under consideration. These differing scenarios need to be considered while implementing the optimizations. Following are the optimizations this paper presents:

1. Moving scripts to footer
2. Loading JS from Google libraries
3. Defer parsing of JavaScript files
4. Remove query strings
5. Lazy load images to improve speed
6. Removes extra Font Awesome styles
7. Remove RSD Link
8. Remove all RSS feed links
9. Load CSS asynchronously
10. Minify all CSS styles
11. Insert all CSS styles inline to the footer

As mentioned earlier this paper also proposes an Analyzer tool that will help analyze the nature of the complexity of a website.

Depending upon this analysis of complexity appropriate optimization mechanisms can be implemented to reduce the ill-effects of complex nature of a website on the webpage load time.

Structure of the paper is as follows:

Section **2** gives a brief overview of the work related to this concept. Section **3** describes the idea of the Analyzer tool and possible structure for the same. Section **4** presents various optimization mechanisms. Section **5** concludes the overall discussion and talks about future scope.

2. Existing System

There has been a lot of work on how to increase the performance of web by considering network traffic, web graphs, page rank etc. But very few efforts have been taken to understand the nature of individual website and its implications in terms of page load time and client performance.

Some of the important work related with this study is as follows:

Reducing Web Page Post back:

Rohit Dhand, in his thesis describes that how Webpage post back calls causes increase in the bandwidth utilisation and thereby page load time of the website [3]. He has proposed the importance of jQuery and Ajax call to reduce the web page post backs.

Rich Client Web Architecture Based on HTML5 (RCWABH5):

CHEN Li-li and LIU Zheng-long, in their thesis proposed the rich client web architecture based on the HTML5 [4]. This architecture is useful to improve client performance by incorporating advance features of HTML5 to replace traditional notions like Plug-ins. This paper also focuses on the importance of increasing the user's experience and client performance without causing increase in the web page load time. This study enables us to find new techniques to balance the trade-off between client performance and users experience.

Web Prefetching to improve performance of websites:

Monti Babulal Pal and Dinesh Chandra Jain; in there thesis proposes the importance of the concept of perfecting in terms of web applications to increase the performance and effectiveness of web services [5]. There work deals with the notion of Web caching to improve the performance of Web based system. Web caching involves keeping Web objects that are likely to be used in the near future in location closer to user. The study considers the

Web caching mechanisms at three levels: client level, proxy level and original server level.

Web Page complexity:

A recent study by Michael Butkiewicz, Harsha V. Madhyastha, and Vyas Sekar has proposed a new concept called Web Page Complexity [2]. This concept deals with the nature of complexity of an individual website in terms of different factors. For quantifying the complexity of a website they have proposed various metrics. Further the concept of web page complexity has been divided into two type viz. Content level complexity metrics and service level complexity. The impact of the webpage complexity on the page load time is another aspect this study deals with. The thesis asserts that the concept of Web Page Complexity is purely new notion in the field of Web and needs to be worked upon to analyze its impact on the page load time and improve client performance by implementing various optimization techniques.

3. Analyzer Tool

As discussed earlier we are proposing an Analyzer As discussed earlier we are proposing an Analyzer tool to analyze the complexity of a website and generate a report regarding the nature of the complexity. The report can thereafter be used to determine which set of optimization mechanisms are best suited to reduce the web page load time. As we have mentioned, complexity of a website can be categorised into two broad type namely 'content level complexity' and 'service level complexity'. Content level complexity represents the number and type of objects requested to load a webpage. Service level complexity represents the number of servers used to load a webpage and nonorigin content on the webpage. So the analyzer tool will analyze a website to generate a report describing the nature of the complexity of the website. This report will be used to determine which kind of complexity is dominant in the website and ultimately affects the page load time. This knowledge can later be used to take the decision regarding the selection of the optimization mechanisms to reduce the webpage load time.

The Analyzer tool works upon various indices to determine the nature of complexity of a website. These indices are NumberOfObject, TypeOfObjet, NumberOfServers, NonOriginContribution. The former two are used to determine the content level complexity while the later two are used to represent service level complexity of a website. By analyzing these indices Analyzer tool will ultimately decide another two indices namely ContetComplexity and ServiceComplexity which denotes the measurement of the two types of complexities. Values of ContentComplexity and ServiceComplexity can thereafter be used to decide the set of optimizations to be carried out in order to reduce the webpage load time.

The general architecture of the system is as shown in the figure 1.

Input to the system is a website whose webpage load time is to be reduced. Analyzer tool accepts the code for the website as an input and by considering all the indices mentioned above determines the nature of the complexity of the website. The output of the analyser tool is passed as an input to the next stage i.e. Optimization.

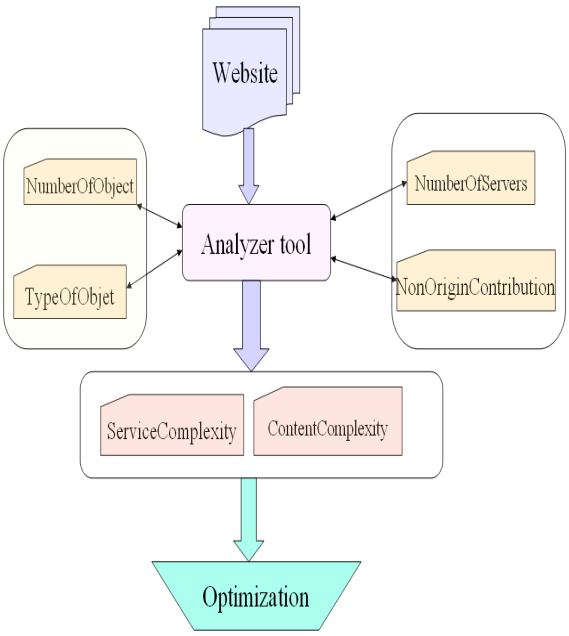

Figure 1: General Architecture of the System

4. Optimization

This section describes some of the optimization mechanisms to reduce web page load time of complex websites. We have selected optimization techniques that may produce larger impacts on the Web page load time of complex structured websites.

1. Moving scripts to footer-Moving Java Script files to the footer of the page will execute all the scripts once the page gets loaded.

2. Loading JS from Google libraries-While using the 3rd party script files such as jquery, It is recommended to

use CDN copy of script file. Google is one of the CDN providers which is used to host popular scripts libraries. These files may already be present in the browser through cache.

3. Defer parsing of JavaScript files-Deferring the execution of JavaScript files can help reduce the page load time of the webpage.

4. Remove query strings-Passing query parameter or version parameter to the included scripts will make unique copy of url which cannot be loaded through the local cache. Each time we need to request the new copy of the script. Removing query string from url will help use same copy of script multiple times.

5. Lazy load images to improve speed Lazy loading of the images helps to load only visible portion of the page. Non-visible portion of the page gets loaded once user starts scrolling.

6. Remove extra Font Awesome styles Font-awesome provides cross-browser compatible libraries to load icon in webpage in different screen resolutions. We can only include the selected icons instead of loading all the icons.

7. Load CSS/JS asynchronously Loading CSS files asynchronously reduces the waiting time of the browser to load the complete page.

8. Minify all CSS styles -It includes the compression of the CSS selectors to reduce the size of the CSS file by removing comments and creating local reference of the image paths.

As we have mentioned earlier, optimization is not a straight forward task, so different websites will need different set of optimizations to achieve better performance.

.

5. Conclusion

Webpage complexity is the new concept that needs to be studied further to get new insights into the issues related to the Web performance. This paper tries to present different optimization mechanisms that can be implemented on a website to reduce the webpage load time by analyzing its complexity with the help of the proposed Analyzer tool. It is also found that optimization is not a straight-forward task and needs to be implemented by considering nature of the complexity of the website under consideration. For this purpose the proposed Analyzer tool can play a significant role by determining the nature of the complexity of a website so that appropriate optimizations can be carried out to mitigate the ill-effects of the complexity on the webpage load time.

References

[1] Keshab Nath, Sourish Dhar and Subhash Basishtha , “Web 1.0 to Web 3.0 - Evolution of the Web and its Various Challenges” 978-1-4799-2995-5/14 2014 IEEE

[2] Michael Butkiewicz, Harsha V. Madhyastha, and Vyas Sekar, “Characterizing Web Page Complexity and Its Impact”, IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 22, NO. 3, JUNE 2014

[3] Rohit Dhand, “Reducing web page post backs through jQuery Ajax calls in a Trust based Framework”, 978-0-7695-4817-3/12 2012 IEEE

[4] CHEN Li-li and LIU Zheng-long, “Design of Rich Client Web Architecture Based on HTML5”, 978-0-7695-4789-3/12 2012 IEEE

[5] Monti Babulal Pal and Dinesh Chandra Jain, “Web Service Enhancement Using Web Pre-Fetching By Applying Markov Model”, 978-1-4799-3070-8/14 2014 IEEE

Omkar B. Sawant received the B.E. degree in Computer Engineering from P.V.P.I.T, Pune, INDIA in 2010 and is pursuing M.E. in Computer Engineering at S.A.O.E , Pune

Sachin P. Godse received the B.E. degree in Computer Engineering from A.V.C.O.E, Sangamner, INDIA in 2004 and M.E. degree in Computer Engineering from S.C.O.E, Pune in 2010. He is currently working as a Professor in S.A.O.E, Pune.