Efficient Techniques for Reducing Web Page Complexity

Anand M

Assistant Professor, Department of Information Science and Engineering, GSSSIETW, Mysuru, Karnataka,

India

ABSTRACT

Users being frustrated due to high page load times. Because page load times are directly effects user satisfaction, providers would like to distinguish if and how the complexity of their Web sites affects the user experience. Although there is an extensive literature on measuring Web graphs, Website popularity, and the nature of Web traffic, there has been little work in understanding how difficult individual Web sites are, and how this complexity impacts the client's experience. We proposed a system that can identify a set of metrics to characterize the complexity of Web sites both at a content level and service level. Also the impact of complexity on the user performance and recommending what control measures should take to reduce the complexity.

Keywords : Browsers, Websites, World Wide Web, Internet.

I. INTRODUCTION

Since many years, Web pages have become expressively more complex. At first web site used to host text and images, To rich media like Flash and Silverlight now the web pages contain several content type, ranging from video to script performed on the client's device. Additionally, a Web site today fetches content from servers hosted by its providers, and also from a range of third-party services such as advertising agencies, content distribution networks (CDNs), and analytics services. In combination, representation of a single Web page today involves fetching several objects with varying characteristics from multiple servers under different administrative domains.

In contrast, the poor effects of slow Web sites are well known. Users will discard or switch a Web Site due to performance issues. According to recent surveys, out of whole two thirds of users encounter slow Web Site. While abundance of anecdotal proof is that a key factor in slowing down Web Page is the increase in Web page complexity. Official studies on this topic have been limited. Most previous work on Web measurement concentrates on characterizing the Web graph [7], study the network footprint of Web traffic [2]-[5], also studying the rate of change of content on the Web [8]. Although these have contributed to a well understanding of Web usage, they do not examine the Web sites themselves.

In this paper we focuses on two broad question. First, we count the complexity, called content level complexity of a Web page by means of a broad spectrum of metrics, and we characterize a Web page by the content in rendering like-the number of objects fetched, the sizes of these objects, and the types of content. Also we study the complexity of Web pages relating to the services they build upon. In addition to these we find number of bytes fetched and non-origin content accounts in place of a significant fraction of the number of objects.

Our second and the main focus is on the time to download and render a Web page. We find that the total number of bytes fetched to render a Web site is the most dominant indicator of client-perceived load times than the number of objects fetched.

II. LITERATURE SURVEY

Michael Butkiewicz, Harsha V. Madhyastha, and VyasSekar [1] are focuse in these paper on finding the gap in understanding how complex individual Web sites are and how this complexity impacts on the usersperformance. Also characterize the Web site both at content level (like, number and size of images) and service level (like, number of servers/origins). It may happen that some categories are more complex than other such as 'News'. Out of hundred 60% of Web sites fetched content from minimum five non-origin sources, and these give more than 35% of the bytes downloaded. In addition, they examine which metrics are most suitable for predicting page render and load times and catch that the number of objects requested is the most important factor. With respect to variability in load times, however, they also find number of servers is the best indicator.

Y. Zhang, H. Zhu, and S. Greenwood [6] discuss about navigability. Navigability has become the axis of website designs. Existing mechanism have problem into two types. The major is to assess and measure a website's navigability in contrast to a set of principles. Another is to evaluate usage data of the Website. A metric methodology to Website navigability measurement is studies in this paper. Objectiveness and the probability of using automated tools to assess extensive websites are advantages of navigability metrics as far the existing valuation and analysis techniques.

Axiomatic assessment

Weyuker's axioms of software complexity have been frequently applied in place of a method to authorizing A logically the measurement of software complexity .In this section, they measure the metrics well-defined in the prior section compared to Weyuker's axioms of software complexity. Weyuker's axioms are established on a number of operators and relations on programs. According to the features of websites these operators and relations must be modified.

M. Lee, R. R. Kompella, and S. Singh[3],in this paper discuss on Cloud-based Web applications driven through new knowledge such as Asynchronous JavaScript and XML (Ajax) place an important load on network operators and creativities to effectively manage traffic. Problem happen is that there is no systematic technique to produce their workloads, notice their network performance today and possess track of the varying trends of these applications. They develop a tool, called AJAXTRACKER, that automatically impersonators a human interface with a cloud application and gathers associated network traces.

Ajax tracker

The main workings of AJAXTRACKER contain an event generator, a Web browser, a packet sniffer, and a traffic shaper. The event generator procedures the bulk of the tool.AJAXTRACKER is agnostic to the select of the Web browser and can work through any browser. The aim is to collect representative traces of a client session; packet sniffer captures the packets proceeding the client machine.

B. Krishnamurthy, C. E. Willis, and Y. Zhang [2], focuses on Content Distribution Network (CDNs). This is a tool to distribute contents just before and users. In content distribution origin server serves some or all the content of web pages. The technique like DNS resending and URL rewriting are balance the load among their servers. After the observation of results some CDNs provide better results as compare to other. In particular network the dramatic growth in the number of distinct server is gives best performance of one CDN company that can be improved between two testing periods. Either in

average or worst case conditions the results shows that the case of DNS in the critical path of resource retrieval is not better than that of server choices related to client response. 2.5. Understanding Online Social Network Usage from a Network Perspective F. Schneider, A. Feldmann, B. Krishnamurthy, and W.Willinger,focues in this paper study of Online Social Networks is discussed. Also, they understand which OSN feature inters and which one keep in consideration of poor users. Additionally the topics like friendship graph and sample crawls are studied on surveys. Extracting clickstreams as of inactively observed network traffic these are the techniques using these they study how users are interact with OSN.

III. PROPOSED SYSTEM

By using the "Characterizing Web Page Complexity and Its Impact" we are finding the complexity of web page. After the complexity has been calculated we can show the analytical reports in text as well as graphical format using graphs.
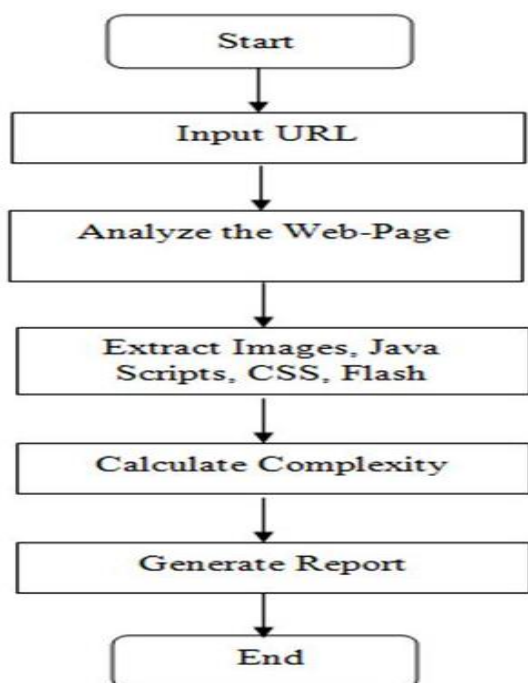

Figure1. Content Flow Architecture

We are starting with entering url after that the page is load. After the loading of page complexity of that

web page can be determine using parameter such as content complexity and service complexity.

HTML5

HTML5 is a new standard for HTML which allows us to build rich and interactive web pages which bring HTML into the world of application development started in the year 2004. HTML moves from simply describing the basics of a text based web for presenting audio, video and animations to enabling offline functionality, geo location and local storage in client side databases.

With the development of HTML5 it has wide range of applications in multimedia direction [3]. It can play audio and video and supports animations from the browser without the need of the proprietary technologies. The features of HTML5 would add up value for web designers and developers.

With the development of HTML5 it has wide range of applications in multimedia direction [3]. It can play audio and video and supports animations from the browser without the need of the proprietary technologies. The features of HTML5 would add up value for web designers and developers.
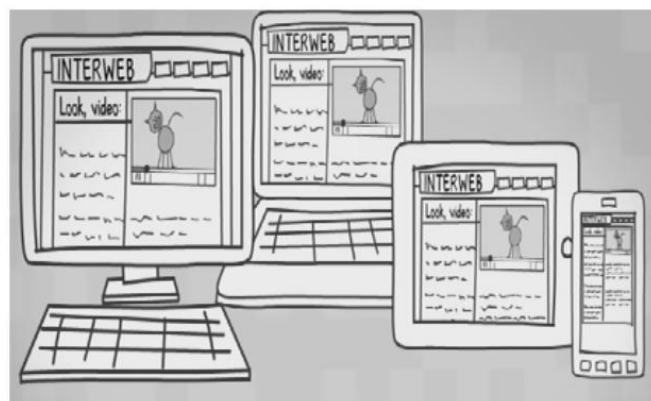

Figure 2. HTML5 cross-platform

HTML5 also support location based services open formats such as Scalable Vector Graphics (SVG), open XML file formats and high quality graphics.

The basic advantage for the developers and browsers is that they would be able to do more without the need of mastering or licensing multiple proprietary

technologies that can develop rich web pages, enhanced forms and web based applications.

HTML5 FEATURES

HTML5 provides new features that include

1. Canvas – 2D/3D Graphics
2. Audio & Video
3. Location based Services
4. Working Offline
5. Web Workers
1. Drag & Drop
2. New Input Types
3. New Elements
4. Form Elements

HTML5 supports cross platform, designed to display web pages on a PC, or a Tablet, a Smartphone, or a Smart TV (Fig-1). HTML5 is been a working draft and some browser designers and websites are already adopting HML5 elements.

IV. CONCLUSION

We will develop an efficient system that will calculate the complexity of the webpage and identify the components of web pages which requires more time to load.

V. REFERENCES

[1] Michael Butkiewicz, Harsha V. Madhyastha, and VyasSekar, "Characterizing Web Page Complexity and Its Impact", VOL. 22, NO. 3, JUNE 2014

[2] B. Krishnamurthy, C. E. Willis, and Y. Zhang, "On the use and performance of content distribution networks" , in Proc. IMW, 2001, pp. 169–182.

[3] M. Lee, R. R. Kompella, and S. Singh, "Active measurement system for high-fidelity characterization of modern cloud applications", in Proc. USENIX Conf. Web Appl., 2010, p. 2.

[4] F. Schneider, A. Feldmann, B. Krishnamurthy, and W.Willinger, "Understanding online social network usage from a network perspective", in Proc. IMC, 2009, pp. 35–48.

[5] F. Schneider, S. Agarwal, T. Alpcan, and A. FeldmanninProc. PAM ,"The new Web:Characterizing AJAX traffic", 2008, pp. 31–40

[6] Y. Zhang, H. Zhu, and S. Greenwood, "Website complexity metrics for measuring navigability," in Proc. Int. Conf. Quality Softw., 2004, pp.172–179

[7] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw., "Graph structure in the Web," vol. 33, no. 1, pp. 309–320, Jun. 2000.

[8] D. Fetterly, M. Manasse, M. Najork, and J. Wiener, "A large-scale study of the evolution of Web pages," in Proc. WWW, 2003, pp.669–678