Framework for retrieval and rendering of information in real time

Prince, Anil Katiyar, National Informatics Centre Punjab

Deepak Kumar, Research Scholar IIT Mandi.

Abstract:

In today's era the information is basic need for everyone, there are lots of work already done in the field of information retrieval and rendering, still there is a some areas where the research are needed to fill the gap. Our approach is related to information retrieval and rendering will try to identify those gaps and present novel solution for addressing the problem in this area. But this entire concept is human centric, that is user have to come up at search interface to find out information. We are proposing a framework where any web site can crawl the content form other sites for their user's need and present the information in own web page. In this concern two type of strategy can be possible one when linked parent site provide the simplified contents to the associated child site (push) by self and second is when linked child site fetch the contents from parent site as it want(pull). The first scenario is well discovered in the form of site feed like RSS and Atom or JavaScript widget such as Facebook "Like Box" to update section of page but in second scenario there is a need of research.

Keywords: crawler, content filtering, rendering.

Introduction:

The representation of information quality, also known as "reality visualization," can be explained as a judgment of the accuracy of specific information in a dataset relative to the whole. The visualization of information quality is an imperative factor in data comparison situations, or even to simply "highlight" regions of low quality information for further study.

Information retrieval and rendering based models have been investigated for a long time. Many contexts have been redesigned as crucial issues in information retrieval and rendering. Content filtering, context features, and context aware are important factors for any information retrieval and rendering system. A highly dynamic search environment and contextual features are not captured at indexing time, nor are they exploited at retrieval time. That retrieval may be inaccurate, since they are formulated while not rendering explicit the features given by users' requirement and objectives. Perception approaches have been incorporated into rendering algorithms in order to optimize rendering computation and produce information according to point of view. This paper represent framework for retrieval and rendering of information in real time.

One approach to retrieving information in context is by re-ranking documents using evidence about informative content, space, and time or by using user derived evidence from document clusters, query histories, profiles, sensor data, clock and cylinders. For example, a set of documents may give a cluster which can be used as description of a context. These types of approaches are not operated, since most common models lack a formal representation of context; for example, it is unknown whether the clustered documents are from the same context. IR models have in the past been defined by ignoring the multiplicity of users, information needs, locations, times, and histories. Indeed, these models are defined by assuming that there is one user, one information need for each query, one location, one time, one

history, and one profile. With such a approach, inexperienced users can begin to discover, explore, and utilize Web data and services. Because data and services are accessed directly through a standalone client and not through a central point of access (e.g., a portal), new content and services can be consumed as soon as they become available. In this way we take advantage of an important sociological force that encourages the production of new Semantic Web content by remaining faithful to the decentralized nature of the Web [5].

User-generated content possesses three important proper-ties: (a) very high level of heterogeneity, which results from its distributed, self-organized creation, (b) high growth rates in terms of volume, and (c) fast evolution, which implies re-activity as well as volatility. This fickleness, in particular, constitutes a salient aspect of user-generated data: the com-munity of users tends to update and extend the information shared on the Web, by continuously adding new content and modifying the existing one. However, not all this information is of the same importance; some parts reflect reality, but others comprise wrong, obsolete or slightly irrelevant content. As more and more users of Web 2.0 systems examine these profiles, they gradually identify the false pieces of information and remove or replace them with more relevant ones [1].

In this paper we present an approach for retrieval and rendering of information. We are proposing a framework where any web site can crawl the content from other sites for their user's need and present the information in own web page. In this concern two type of strategy can be possible one when associated parent site self-provide the updated contents to the associated child site (push)and second is when associated child site fetch the contents from parent site as it want(pull). The first scenario is well discovered in the form of site feed like RSS and Atom or JavaScript widget such as Facebook "Like Box" to update section of page but in second scenario there is a need of research. The aim to create more effective way of presenting information through the immersive environment in which users navigate through special features, structures and scenarios.

Framework:

Search engines have been a mechanism to organize and index information available on the web and bring up the most relevant results for a particular query. This mechanism has constantly evolved with the changing Internet usage patterns and growth in content and media types. One question that we often get asked is, what is the future of search? There are no easy answers. However, it is important to look at the progression and usage trends to determine how search engines of the future would work. Yes, while we still use traditional search engines, we're relying more and more on these newer technologies to guide us to both specific content that we're looking for as well as content that we don't even *know* we're looking for (as in the case of suggestive search). It is my belief that the future of web development lies in this field, which we call *non-traditional search i.e.* retrieval and rendering for sites user base. Let us consider a scenario in which student come up on his college site and want to get the latest circular updated by college as well as the university. In today's situation if the university's site having the provision of latest feed than site of college can bring it by updating page through RSS reader and publishing mechanism. But in the case when university's site does not have any provision of feeds student have to visit both the college and university site separately.

Our solution will bridge the gap of previous discussed problem. In the given research work proposal we are trying to build the **"framework for retrieval and rendering of information in real time"** the protocol we are developing to providing communication between information provider site and receiver site is **IPCUP** dubbed as **"Inter page contents update protocol"**. The working of proposed framework can be discussed with the figure 1. As shown in figure the proposed framework will have a crawler to crawl the contents from provider site, content filtering module, categorization module

and real time rendering module etc. The basic mechanism of said protocol to authorize receiver and
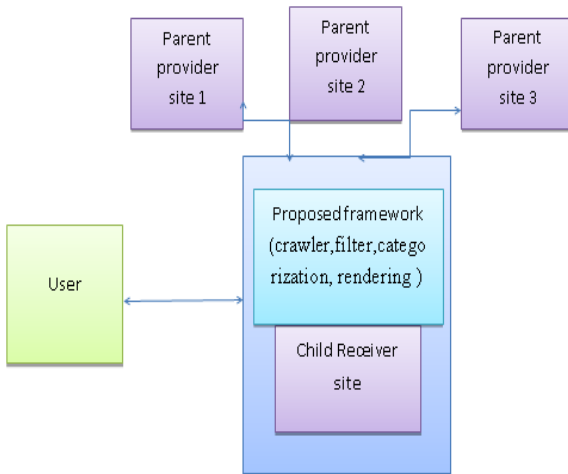

Figure 1.

control the section of page crawling from provider sites for real time rendering on receiver site based on provision. The overall advantage of this framework is one stop information aggregation according to child site.

The content of different sites is filtered at child server. At filtration step, we extract only similar data in different site. There are different types of filtering algorithms are available, that can be used for content filtering. After filtration process, then categorized the content according to scenario/requirement. During categorization process, compile similar content and avoid repeated data. So that data replication can avoid. After categorization, rendering process will start. In rendering process, we apply some predefined graphics, style functions and methods on categorized data. This mechanism has constantly evolved with the changing usage patterns and growth in content and media types.

Result and Future Work

To verify our work, the user study to be conducted will comprise a series of tasks that are focused on the user's ability to identify creatively rendered

information as uncertain or unspecified lower quality. The user will be presented along with a task description and task details, and then asked to complete the task several times with variations in categorization, orientation, and rendering method. The retrieval and rendering process used are not based on actual data. Effectiveness of the means of uncertainty representation will be analyzed according to the accuracy of the tasks performed, retention of information to indicate higher or lower cognitive loads required for task performance, and through subjective user responses to the methods of information rendering technique. After an initial study to identify the most effective methods of representation within artistic rendering techniques, a scenario-based study will be conducted to further evaluate the effectiveness of the techniques on a cognitive level.

Conclusion

Our approach is related to information retrieval and rendering will try to identify those gaps and present novel solution for addressing the problem in this area. In current scenario of search engine is to crawler the content from web and provides the significant result against user query. But this entire concept is human centric, that is user have to come up at search interface to find out information.

Acknowledgments

Reference

[1] J. Surowiecki. The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes busi-ness, economies, societies, and nations. Random House, Inc.,2004.
[2] Kuhlthau, C. C. Seeking Meaning: A Process Approach to Library and Information Services. Ablex, Norwood, NJ, 1993.
[3] Kumar,N.Andbenbasat, "The effect of relationship encoding, task, type, and complexity on information representation: An empirical evaluation

of 2D and 3D line graphs.", MIS Quart. pp 255–281, 2004.

[4] Storey, V., Burton-Jones, A. Sugumaran, V. Andpurao, S. Conquer: A methodology for context-aware query processing on the World Wide Web.Inf. Syst. Res. 19, 1,p.no. 3–25, 2008.

[5] C. Bajaj, V. Pascucci, and D.R. Schikore. Fast is contouring for improved inter-activity. In Proceedings of the 1996 symposium on Volume visualization. ACM, October 1996.

[6] Dennis Quan, David R.Karger. How to make a semantic web browser, ACM, 17-24, May-2004.

Biographies

DEEPAK KUMAR, received the B.Tech degree in Information Technology from the BIET Jhansi (Uttar Pradesh Technical University, Lucknow) in 2008, the M.Tech degree in Human Computer Interface from IIIT Allahabad in 2011. Currently he is P.hd Scholar at IIT Mandi. He can be reached at deepaktyagi12@gmail.com

PRINCE, received the B.Tech degree in Information Technology from the LIT (Punjab Technical University, Jalandhar), Punjab, in 2007, the M.E. degree in Computer Science Engineering from the PEC University of Technology, Chandigarh, in 2010. Currently, He is an District Informatics Associate, National Informatics Centre Faridkot. Prince can be reached at er.princegoel@gmail.com

ANIL KATIYAR, received the B.Tech degree in Computer Science and Engineering from BIET Jhansi (Uttar Pradesh Technical University, Lucknow) in 2008. Currently, He is an District Informatics Officer, National Informatics Centre Faridkot. Mr. Anil Katiyar can be reached at aanilkatiyar@gmail.com