Proceedings of Student-Faculty Research Day, CSIS, Pace University, May 2nd, 2008

A Proposed Framework for Identifying Data Structure Standards for Search Results

Mohib Fanek (*Mohib@hotmail.com)*
Seidenberg School of CSIS, Pace University, NY

Abstract

Computer technology is moving at a very fast pace. It is not unheard of for a personal computer to have a five hundred gigabyte hard drive. To put that into perspective, that equates 6,500 word documents (avg. size 750 Kb), 1,7000 mp3s (avg. size of 3Mb), or 3,000 images (avg. size 1.5 Mb). The World Wide Web contains a much wider array of information. It also provides a "seamless integration of heterogeneous data from distributed sources, letting agents (human users or automated programs) perform sophisticated and detailed analyses of the data" [4]. I propose a framework for a standard data structure for search results using XML. The data structure can be integrated into other applications aside from web search such as database queries and desktop searches.

1. Introduction

A visually enhanced display of search results creates a sense of presence which has been shown to "impact a user's ability to perform a task and, therefore, that insight into presence offers a potential payoff in terms of task performance" [1]. The advancement of the methods used for presenting search results on the web "… contrasts sharply with the advances in design of search techniques that allow indexing large volumes of information and efficiently executing keyword-based search" [3]. Many methods for presenting search results exist but they utilize different structure for returning the data. If a standard structure for storing the search results to any query method is used, the end user will have many options for displaying those search results. The options range from a flat 2D list (used today on the web) to a three dimensional hyperbolic tree (also known as hypertree).

2. Desktop search

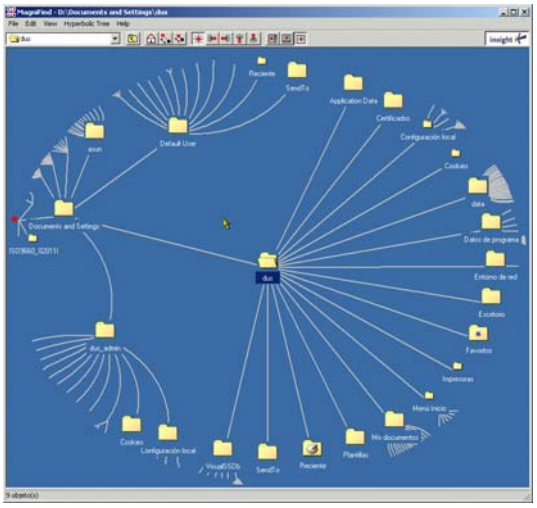
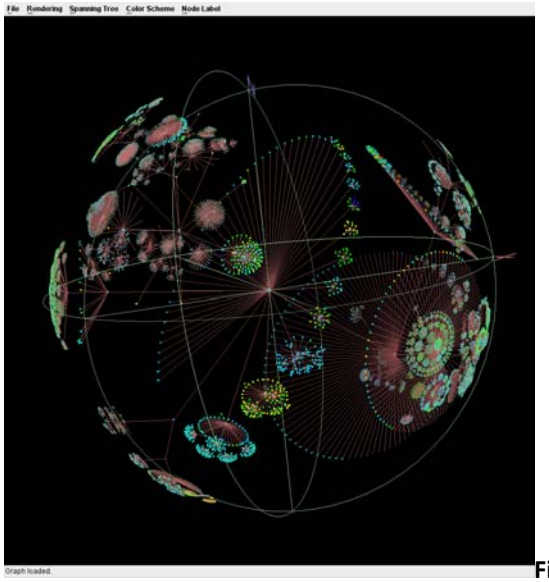


Figure 1. Hyperbolic geometry: 2D representation of the directory structure of a PC [5].



Fig

Figure 2. Hyperbolic geometry: Spectacular view of a complex directory structure in 3D hyperbolic geometry of a complex tree of directories [6].

3. Web Search

Search engines today focus on returning a narrow search result in a short amount of time. Not enough

emphasis is placed on the display of the data; specifically, the correlation between data on distributed sources. Already existing standard data structures the present web search results in 3D space such as Perspective Wall, CHEOPS, Cone tree, Information cube and future applications can be adapted to the visualization of web spaces [2]. For this to happen, the language and expressions used to describe the search results should be standardized so that these different visual paradigms are applicable. For example, a search may return fifty thousand results which, if displayed in a flat list with hyper links, would take 2,500 pages displaying twenty results per page. In a hypertree, the pages can be grouped based on common tags between the pages and displayed hierarchically so all the results are presented at the same time, eliminating page navigation.

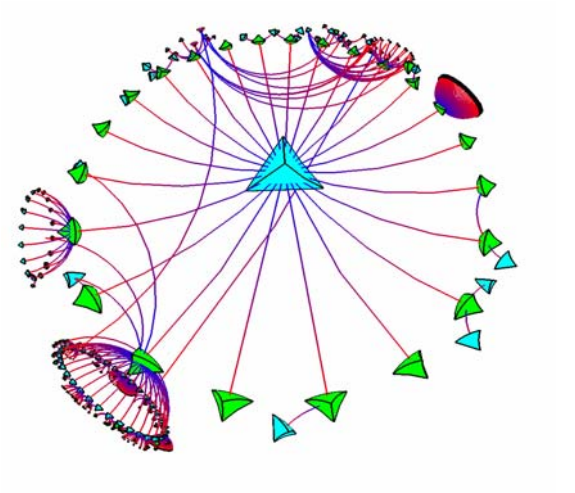


Figure 3. Hyperbolic space (the projective model) [7].

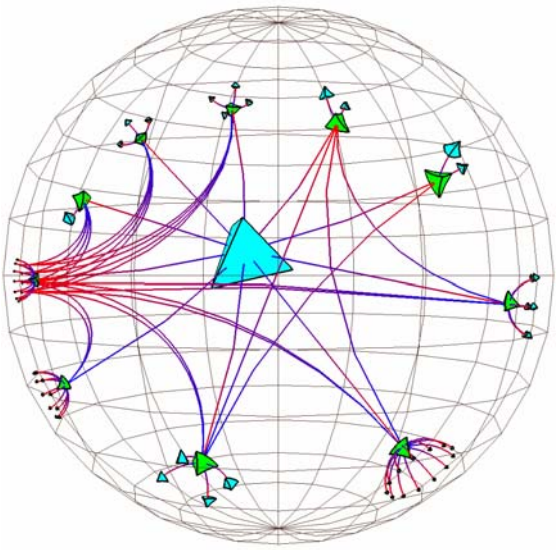


Figure 4: Hyperbolic conformal model of hyperbolic space [8].

Applications and Benefits of a Standardized Data Structure

A standardized search result structure represents a layer between the data and the UI (User Interface). Take the UI choices available for users today are:

- Desktop search:
  - File / Folder List
  - Flat List
- Online Search
  - Flat web page with links and suggestions

With a standardized search result data structure, the UI choices available for desktop search become available for Online search and vice versa.

References

1. Y. Christine and H. Odette, "The Relationship Between Presence and Performance in Virtual Environments: Results of a VERTS Study," 2003, pp. 277.
2. T. Ozgur and S. Ramesh, "Visualization of web spaces: state of the art and future directions," *SIGMIS Database*, vol. 38, no. 3, 2007, pp. 51-81; DOI http://doi.acm.org/10.1145/1278253.1278260.
3. S. Rahul and H. Ya-Wen, "Analysis of usage patterns in experiential multiple perspective web search," *Book Analysis of usage patterns in experiential multiple perspective web search*, Series Analysis of usage patterns in experiential multiple perspective web search, ed., Editor ed.^eds., ACM, 2007, pp.
4. M. Martin, et al., "Retrieving and Semantically Integrating Heterogeneous Data from the Web," *Book Retrieving and Semantically Integrating Heterogeneous Data from the Web*, Series Retrieving and Semantically Integrating Heterogeneous Data from the Web 19, ed., Editor ed.^eds., 2004, pp. 72-79.
5. http://www.infovis.net/printMag.php?num=173&lang=2

6. http://www.infovis.net/printMag.php?num=173&lang=2

7. http://www-graphics.stanford.edu/papers/webviz/

8. http://www-graphics.stanford.edu/papers/webviz/