Journal of Web Engineering, Vol. 3, No.3&4 (2004) 281-296

WEB-BASED VISUALIZATION INTERFACE TESTING: SIMILARITY JUDGMENTS

SHERRY KOSHMAN

School of Information Sciences, University of Pittsburgh
135 N. Bellefield Ave.
Pittsburgh, Pennsylvania 15260 U.S.A.

e-mail: skoshman@sis.pitt.edu

Received July 19, 2004
Revised October 16, 2004

TouchGraph is a Web-based ranked similarity list browser that visualizes the relationship between the query and resulting item set as a graph. TouchGraph provides visual analogs to Amazon's recommendation feature based on item similarity and Google's "similar to" pages. TouchGraph may be able to assist diverse Web users, who have varying levels of knowledge on search topics, to visually select similar items to their query. To examine this assumption further, this investigation asks: *what are the effects of topic knowledge level on the similarity judgments generated by the users in comparison to the visualized system depictions?* Seventeen participants were asked to use TouchGraph for similarity matching of search output to the query and their results were compared to the items shown as most similar to the query by the visualization. The results showed that participants rated their topic knowledge level quite low for most tasks, there was a high degree of participant-system item selection overlap, and a statistically significant relationship was found between knowledge level and node use for half of the tasks. The subjective satisfaction data were positive for the TouchGraph interface. The findings suggest that the TouchGraph visualization has the potential to enhance Web search effectiveness. This study aids in understanding better system design issues in regard to visualization-based tools for Web information retrieval.

Keywords: information visualization, user testing, web search
Communicated by: A Spink & C Watters

Introduction

In information retrieval (IR), a visual approach to displaying queries and data sets offers the potential to see new relationships among data elements that would otherwise remain invisible to the user. Unlike scientific or medical visualization work, which is often modeled on real-world or physical entities, information visualization techniques synthesize abstract data to create visually meaningful patterns and representations of conceptual interrelationships.

Information visualization's foundation is rooted in taking advantage of human visual processing capabilities and responses to visual stimuli [1]. From a systems perspective, visualization-based IR can surpass limitations of text-based systems by presenting users with the opportunity to see their queries in context of the resulting data set, and to see patterns of related items within the data set itself.

From a user's perspective, the World Wide Web is a daunting information retrieval source. The Web's contents are large, volatile, and often structurally ambiguous. Users have several search engine options for query entry and most Web search requests return a deluge of information in the form of text-based lists. Unlike many traditional online IR environments, human intermediaries are not required for formulating searches or evaluating results. Users with varying levels of knowledge on a particular subject can search in any subject domain and they need to make judgments regarding the information they retrieve.

Part of the formula for enhancing Web search effectiveness is based on strategies employed to deal with Web search results. Techniques for helping Web users with their search outcome are generally associated with text-based approaches. For example, a statistical analysis of query terms used to rank items in a retrieved item list is a prevalent method employed to help users select useful information [2]. Human-mediated directory listings may be compiled and classified on the basis of topical relationships to help users make choices.

Web users may be asked to identify links they think are most relevant to their search, or to look at automatically generated "similar page" items in search engine results, such as Google's[a] [3]. Users can then select items on the basis of how closely they match the query or how similar they are to their query. Approaches to system-based similarity searching for Web-based information have been explored [4].

A fundamental function of visualization tools is that they effectively present similarities through methods such as grouping, connecting or placing document icons in relation to each other and/or to the query icons on the screen. This investigation focuses on users' visualization-based similarity selection of Web search results using a ranked similarity list browser, TouchGraph. It is one of the first user studies conducted on a Web-based implementation for visualized information retrieval.

Related Work

Visualization tools currently available for Web information retrieval include systems such as Kartoo[b], Grokker[c], Webbrain[d] and Inxight[e]. Kartoo is a meta search engine that uses cartographic principles to display search output as document maps. Grokker uses colorful nested circles (or squares) to cluster search output and individual items in each category. Webbrain presents a graph-like representation with a focus+context approach to visualize the query as the main focus node with lines emanating to surrounding related nodes. The Inxight Star Tree is based on the hyperbolic tree browser, which visualizes large information hierarchies [5]. It uses the focus+context technique along with colored nodes and links, and features such as spotlighting to visualize Web site information may be applied.

Examples of specialized Web collection visualization include systems such as Anacubis[f] and Stamen.com's view of Google News[g]. Anacubis presents an example of graph-based visualization for business information which uses a colorful hub and spoke model to display a company icon as the central node and uses colored links and icons to visualize its competitors, Web site information, and personnel. Google News is visualized by Stamen.com via a patchwork quilt representation of proper nouns in the daily news. The color and size of the patch represents the item's increased or decreased coverage over time and the extent to which it is covered.

These systems present relatively novel approaches to viewing Web content, and how they impact Web search effectiveness is a new area of exploration. In [6] fifteen users were tested with textual and visualized clustering interfaces using Grokker, Grokker text, and Vivisimo[h]. Vivisimo is a

---

[a] www.google.com
[b] www.kartoo.com
[c] www.groxis.com
[d] www.webbrain.xom
[e] www.inxight.com
[f] www.anacubis.com
[g] www.stamen.com
[h] www.vivisimo.com

textual clustering search engine that utilizes the tree metaphor like Windows Explorer folders to expand a cluster hierarchy into subclusters of information. The researchers used factual information tasks and their results showed little statistical significance for the objective measures among the three interfaces, however participants strongly favored textual interfaces in their subjective satisfaction responses.

Examples of prototype systems used for visualizing Web information include WebVIBE, HuddleSearch, and MetaCrystal. WebVIBE uses a magnet metaphor to represent keywords in a display and the resulting document set is positioned in relationship to the strength of the attraction ("magnetism") between the keyword and document [7]. A usability study with a simplified WebVIBE interface was conducted with participants recruited from the academic community. The participants were not trained on the system and the findings showed that participants found WebVIBE's document animation feature helpful in differentiating document items in various regions of the display for certain tasks.

HuddleSearch uses a clustering algorithm to organize search result sets and a visualization tool to display the contents of a cluster [8]. Researchers recruited 16 users from a university setting to test HuddleSearch against a traditional text-based system. Participants were given eight tasks and used the TREC (Text REtrieval Conference) .GOV collection. Task completion overall was higher and task timings decreased over use with the experimental system. The results were positive for the experimental system in the subject satisfaction reports and differences between the two systems were statistically significant in regard to task and task completion time.

MetaCrystal's approach allows users to evaluate documents returned by several search engines or queries and the visualization is based on the extent of similarity and overlap between the resulting items [9]. The "Category View" displays a document count of items returned by multiple search engines in positioned circles. The "Cluster Bulls Eye View" shows the entire document set and visualizes overlap through proximity and various iconic coding. The model is in the development phase and is not reported to have been tested with users.

The current study lays the foundation for user testing Web-based information retrieval visualization systems by adapting some methodological components found in user studies with prototype visualization systems ([10], [8], [11]). Specifically, the use of structured query tasks, interactive system training, and questionnaires for data collection were adapted to this study.

TouchGraph

TouchGraph[i] is a Java based open source package that provides users with an interface to submit a query and browse visually the resulting set of items that are related to that query. Examples of the TouchGraph interface include a visual analog to Google's *similar to* feature using Google's API (Application Programming Interface) and the *recommendation* feature found at Amazon.com[j].

Google's *similar to* feature offers users the opportunity to find Web pages similar to those identified by the query. The algorithm that powers this "similarity" feature provides the user with additional listings that are ranked and are similar to the identified document (Haveliwala et al. 2002). Amazon provides real time recommendations to online consumers who shop for items from their database. The algorithm used is an item-to-item collaborative filtering model, which compares items that the customer has bought or rated with similar items, then combining these into a recommendation list (Linden et al. 2003).

[i] www.touchgraph.com
[j] www.amazon.com

The TouchGraph visualization uses a focus+context and spring-layout techniques to present the Google and Amazon similarity-based data. Focus+context refers to a visualization that presents the user with an information overview while being able to examine specific details. Hence it provides "context" and "focus" at the same time for the user [12]. TouchGraph allows the user to examine a specific node while seeing the graph overview. The user can also zoom out or in to the graph to view more or less items simultaneously.

Embedded spring theory refers to a process of characterizing the similarity and dissimilarity of documents based on the spring tension forces. Imagine pulling apart a large spring, a point will be reached where the spring and the force pulling it will reach an equilibrium. Conversely, if a spring is pushed toward the center, the pushing force and the spring's restorative force reach an equilibrium. This feature is evident in TouchGraph when a user drags or repositions a node, this action creates a spring-like force that results in node positioning equilibrium.

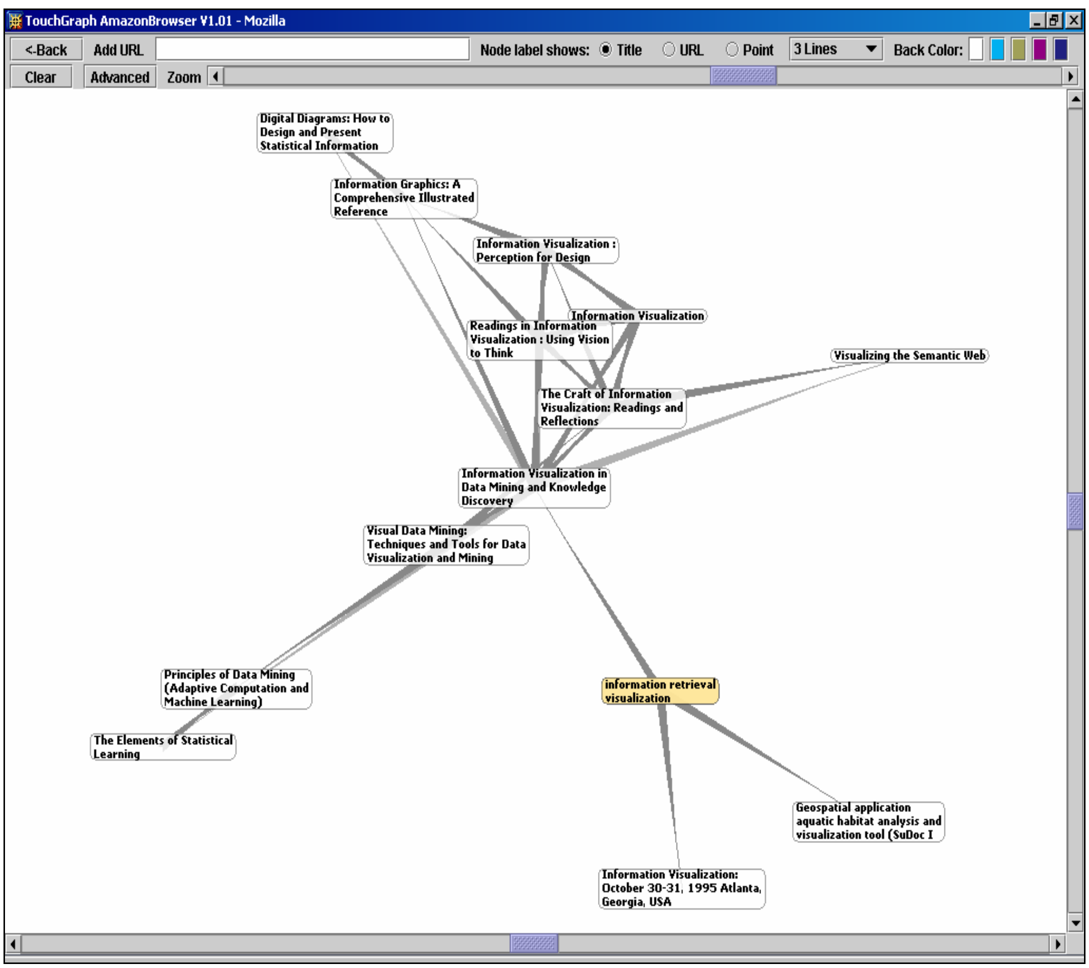

Fig. 1. TouchGraph Layout using the Amazon Implementation for the Query "Information Retrieval Visualization".

To begin a search, the user types a word-based query (Amazon) or a Uniform Resource Locater - URL (Google). The graph places the user's query in the center and features surrounding links to nodes of related retrieved items. (See Figure 1). If the user double-clicks on one of the "nodes", then additional related items are retrieved and graphed around the selected node. The node is marked with a green "C" if there are no further links. Users can obtain additional information regarding an item by moving the mouse over a node and clicking on an "info button". This brings up a small pop-up window with textual information possibly containing additional links that the user can navigate.

TouchGraph's "radius" feature is used to denote the degree of similarity. A radius of 1 means that the surrounding nodes featured will be directly similar to the selected node. A radius of 10 means that items are less similar and are related to other items by a maximum of 10 similarity links. TouchGraph contains a set of visual controls for features such as background color and selecting node label options. Nodes may also be expanded, hidden or collapsed. Single nodes may be viewed, which means they are only related to one other node.

TouchGraph has been cited by [13] as having an important role in eventually visualizing networked Web information to observe Web site growth and development. TouchGraph provides a useful environment in this experimental context since it represents a future direction for user-oriented Web visualization tools.

Research Questions

The primary research question explored here is: *what are the effects of topic knowledge level on the similarity judgments generated by the users in comparison to the visualized system depictions?* Related questions include: a) How diverse is the topic knowledge among participants?, b) What is the impact of topic knowledge on the use of system features to select items?, c) How frequently are the system's features used to make similarity judgments?, d) How much do the user and system sets overlap?, and e) How satisfied were the participants with the interface for item selection?

The level of user topic knowledge is the independent variable and the dependent variable is the similarity judgment. User topic knowledge level is operationalized by a self-defined Likert type scale and defined as the extent of familiarity with a given topic. Similarity refers to user-identified items that relate conceptually to the query presented to the system.

The visualized system is operationalized by using TouchGraph (TG), a Ranked-Similarity-List-Browser System for Web searching. System-generated similarity is defined as the items visualized in the first level of graph nodes surrounding the query. Similarity judgment is measured by:

1. Comparing the number of similar items (C_1) that users select (S) to the similar items depicted by the visualization graph (V) indicated by the first node level. $C_1 = S \cap V$
2. Determining the number of similar items (C_2) that the users selected that are not in the first node level of the graph. $C_2 = S - V$
3. Determining the number of similar items (C_3) that are in the first node level of the graph that the users did not identify as similar. $C_3 = V - S$

It is hypothesized that users with a greater topical knowledge will identify a greater proportion of items related to the query that will match a higher proportion of the graph-generated items. They will click on fewer graph nodes, pop-up windows, and consult fewer links since it is assumed that they need less detail to make their similarity judgment.

Research Design

Participants were exposed to all conditions of the study and a within-subjects design was used. Three of Shneiderman's human factors goals [14] including system learning time, system feature retention, and subjective satisfaction were used as guidelines for developing the methodology, but were not empirically measured.

Data collection

Seventeen volunteers from a graduate level Online Retrieval class at the School of Information Sciences participated in the study as part of an in-class assignment. The majority of the participants were female (70%). All participants were pursuing a graduate level education and had taken at least one information retrieval class. The majority of participants (82.3%) used the Web many times a day, 11.3% used the Web once a day, and one participant reported using the Web a couple times a week. Most participants (66.6%) rarely used visualization-based tools while searching the Web and 33.3% never used these types of tools. The group had no prior experience with TouchGraph.

Materials include a familiarity time worksheet to record observations regarding TouchGraph, a user profile questionnaire to gather basic demographic data, task assignment sheets, and a post-search questionnaire. These items were pilot tested and refined prior to the study.

The study was conducted in two sessions: the first was an exploratory learning session (familiarity time), and the second session included questionnaire distribution, training, and task assignments. In the first session with TouchGraph, participants were given a brief introduction to the system and asked to record their observations while interacting with the interface. Participants were allowed to ask the investigator questions during the one-hour familiarity time session.

In the second session one week later, participants were given their profile questionnaires. They were trained by the investigator on using TouchGraph and their task worksheets were distributed. Randomly assigned tasks were conducted with both the Amazon and Google implementations of the TouchGraph browser. Participants were given a worksheet for each task in which they were instructed to: 1) read the task, 2) rate their topic knowledge level on the topic using a Likert-type scale, 3) enter the provided query on TouchGraph, 4) select similar items from the display, 5) write down the title of items on the worksheet, 6) record the frequency with which they used TouchGraph's nodes, "info" buttons, Web sites in the pop-up window, and 7) list additional features they used.

Table 1. Query Tasks

Amazon
Perennials (PER)
Japanese Cooking (JPC)
DSL
London Travel Guide (LON)
Google
PBSkids.org (PBS)
www.brintannica.com (ENC)

A list of query terms for the task topics is provided in Table 1. There were four tasks formulated for the Amazon implementation using term-based queries and two for Google using URLs. Query task construction was based on Web query log research which shows that the majority of user Web queries tended to be shorter in length, most frequently one or two term queries [15]. Data were collected from six tasks in which participants were asked to retrieve five items for each task from the display that were similar to the query input. Task completion was defined when five items were identified by the user.

After finishing all six tasks, participants filled in a table with their subject knowledge ranking for the task, the number of similar items selected, and their usage of pop-up windows, Web sites and nodes. They also completed a post-search questionnaire to gather information on their overall searching experience and subjective satisfaction with TouchGraph.

Data analysis

Qualitative analyses were conducted on the open-ended responses. Quantitative and statistical analyses were performed on the data using the MS Excel statistical module and Instat software.

Familiarity time

The results from the participants' exploratory session with TouchGraph show that they were primarily concerned with usability issues such as the presentation of TG features to complete their tasks, understanding the document display, and the methods of entering a query. Half of the participants (52.9%) described positive features of the TouchGraph interface such as the expanding nodes option, the "info" button, and the overall node grouping display.

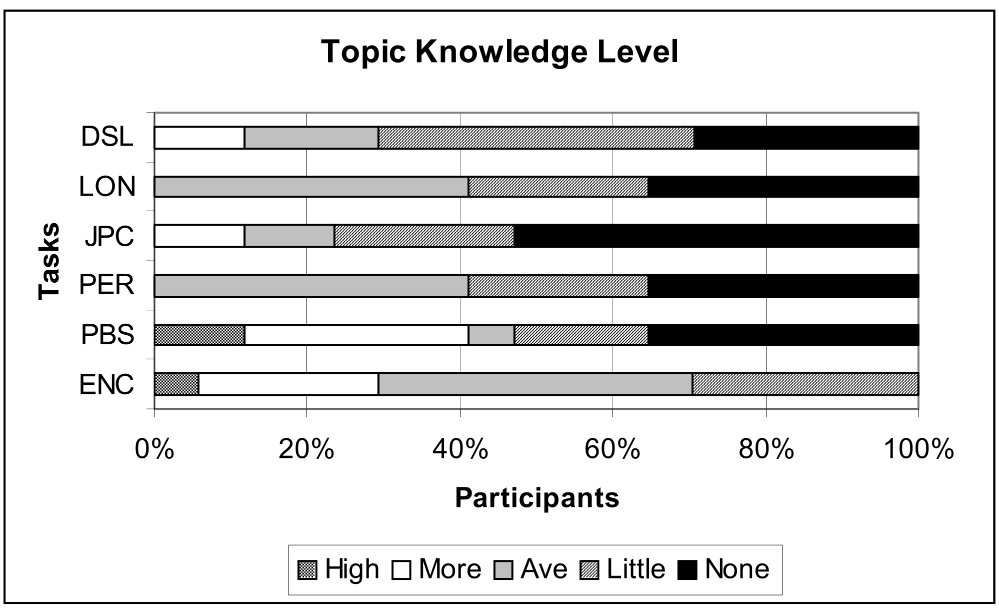

Fig. 2. Level of Knowledge on Task Topic.

Many participants (47%) were uncertain about the utility of entering a URL in the Google Browser implementation since it required the user to know a specific piece of data to enter a search. Also, 41% expressed uncertainty regarding the relationship among items on the screen. Other comments cited problems with overlapping nodes (29%), and the use of color (23.5%).

Diversity of topic knowledge level

At the beginning of the task assignment participants were asked to self-rate their level of knowledge regarding the task topic. Figure 2 presents the participants' rating of knowledge level according to a Likert-type scale with five categories: a) a high amount of knowledge about this topic; b) more than average knowledge about this topic; c) average/basic knowledge about this topic; d) a little knowledge about this topic and e) no knowledge about this topic.

Many participants rated their topic knowledge in the "little knowledge" or "no knowledge" categories. One interesting observation is that the high knowledge level category was only selected for the Google-based tasks.

Participant and system set overlap

The overlapping similarity sets between the user generated and system presented items, and the number of inclusive item overlap among the users were calculated. The participants' selected five similar items to the query from TouchGraph's visualization of the search results and task completion rate was 100%. These items were compared to the browser-generated items that were most similar to the query. This list was compiled by utilizing TouchGraph's advanced Radius feature which displays those nodes that are directly similar to the central node when the parameter is set to one.

The data show that there was quite a high degree of overlap between the user and system sets ($C_1 = S \cap V$). Figure 3 presents the participant and system set intersection, which shows little variation among the tasks. The highest value is for the Google PBS task in which 85% of participants generated item sets that overlapped with the system set. The lowest percentage (67%) is found for the Amazon Japanese cooking task.

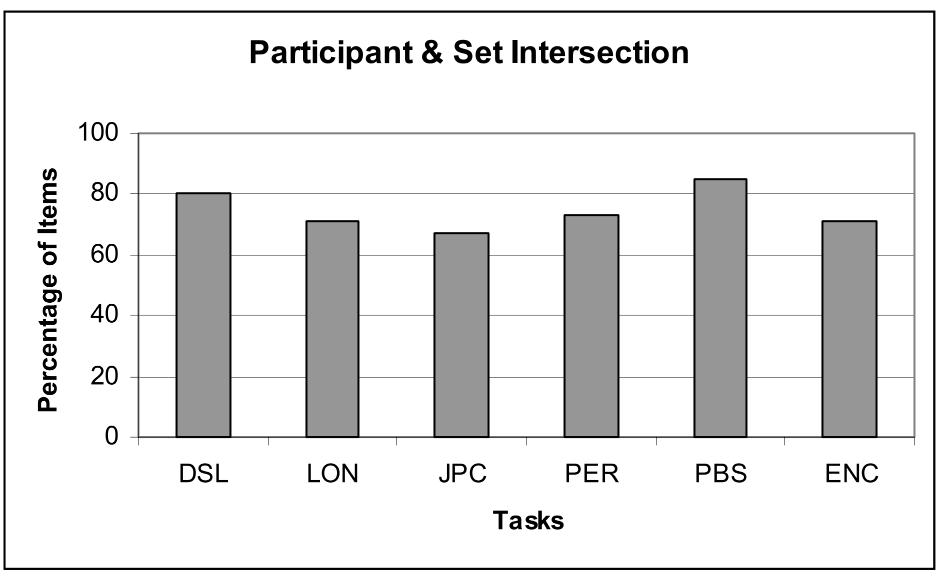

Fig. 3. Participant and System Item Set Overlap

Frequency of system feature use

The participants were asked to select one of five categories to determine the number of nodes used and the number of times they accessed TouchGraph's "info" buttons and Web sites. The categories were numeric ranges representing no use to high frequency and included: a) 0; b) 1-10; c) 11-20; d) 21-30; and e) 31+.

Participants largely selected the 1-10 category for the use of nodes and "info" buttons (Figure 4). The number of nodes in the 1-10 category was selected by a majority of participants (58.8%) for the two Google-based tasks.

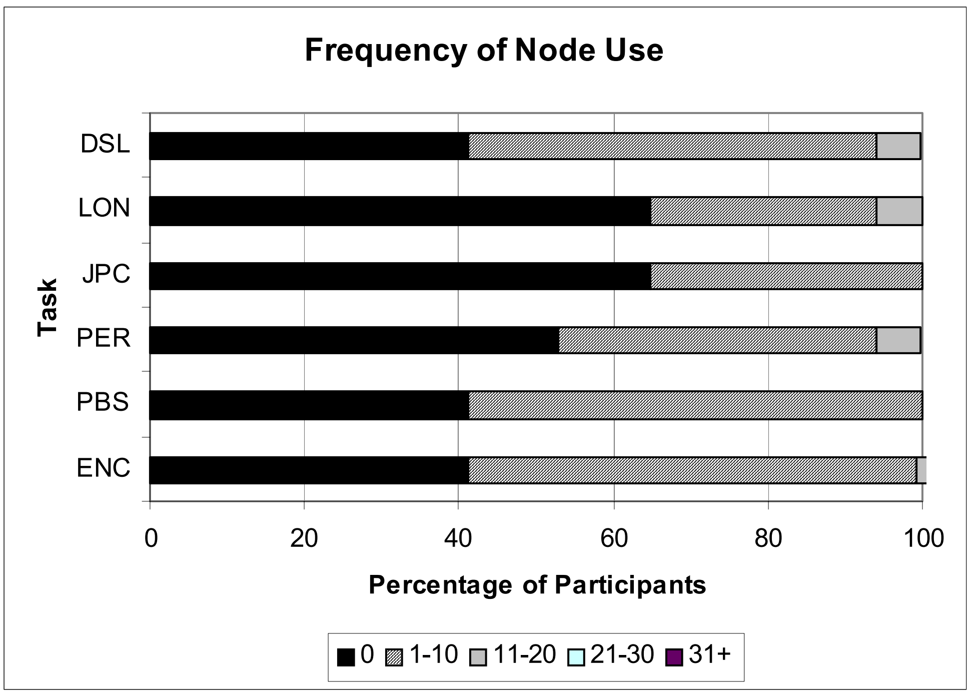

Fig.4. Participant Rating of Node Use Frequency

In regard to "info" button use, the highest percentage of participants (76.6%) selected the 1-10 category for the Google Encyclopedia Britannica task (Figure 5). Similarly, the highest percentage of participants (23.5%) selected the 11-20 category for the PBS Google task. A small percentage of participants (5.8%) selected the 21-30 category for three of the six tasks. The majority of participants (58.8%-88.2%) never explored Web site information available in the pop-up window for the tasks.

The participants' open-ended responses regarding the use of additional features showed that zooming and scrolling were predominantly used for controlling the display. Specific features listed for manipulating nodes included: 1) hide/expand nodes; 2) changing the node labels from displaying three lines of text to one line of text; and 3) using line color to determine node relationships. Only one participant used the advanced radius feature.

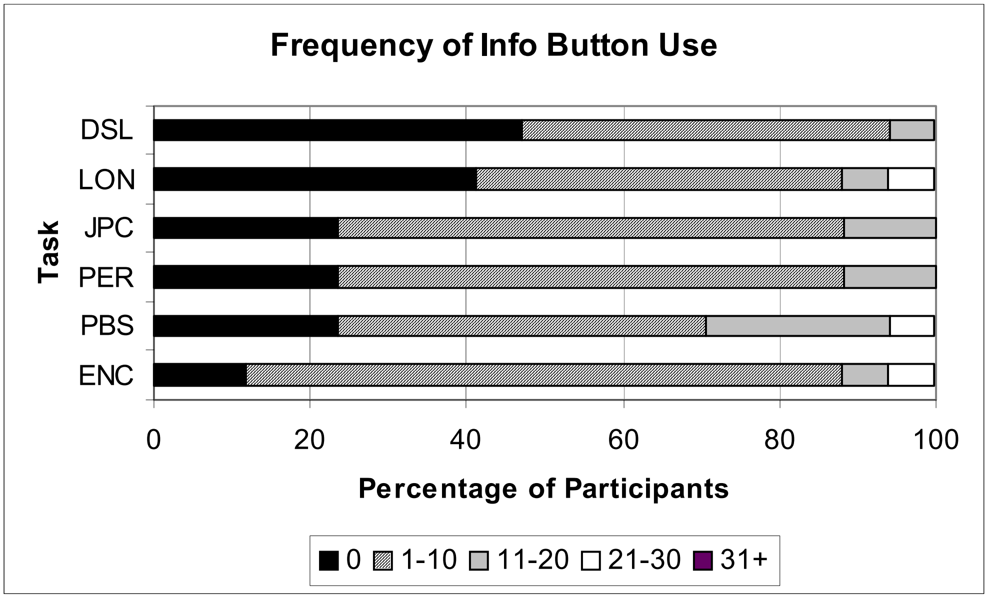

Fig.5. Participant Rating of Info Button Use

Impact of knowledge level on system feature use

A regression test was performed to determine if there was any predictive value of task knowledge level on set intersection, the use of nodes, information buttons, and Web sites used to complete the task. Table 2 shows that the only statistically significant relationship was between task knowledge level and node use ($p<0.001$).

Table 2: Regression Test on Knowledge Level Variable

Knowledge Level	r^2	**p-value**
Set Intersection	0.0009	<0.921
Node Use	0.104	<0.001
Information Buttons	0.002	<0.633
Web Sites	0.007	<0.401

Table 3: Regression Test Per Task

Knowledge Level & Nodes	r^2	**p-value**
Ency Britannica (ENC)	0.395	<0.0069
Perennials (PER)	0.378	<0.0085
London Travel (LON)	0.468	<0.0035

Task knowledge level and node use was examined in more detail among the six tasks. The regression test showed that three of the six tasks showed a statistically significant relationship between these variables (Table 3).

Subjective Satisfaction

The post-search questionnaire measured participants' subjective satisfaction in using TouchGraph. Questions relating to the frequency in which they were able or were not able to locate items related to the query were asked. The participants then rated task difficulty level and the understandability of the TouchGraph display.

Part of subject satisfaction is related to the participants' perceptions of system control and being able to use it successfully to complete tasks. This is significant in the Web searching context as search hits may often be overwhelming to users. The majority of participants reported that they were able to find items that related to the query using TouchGraph Amazon or Google frequently or all of the time (Figure 6).

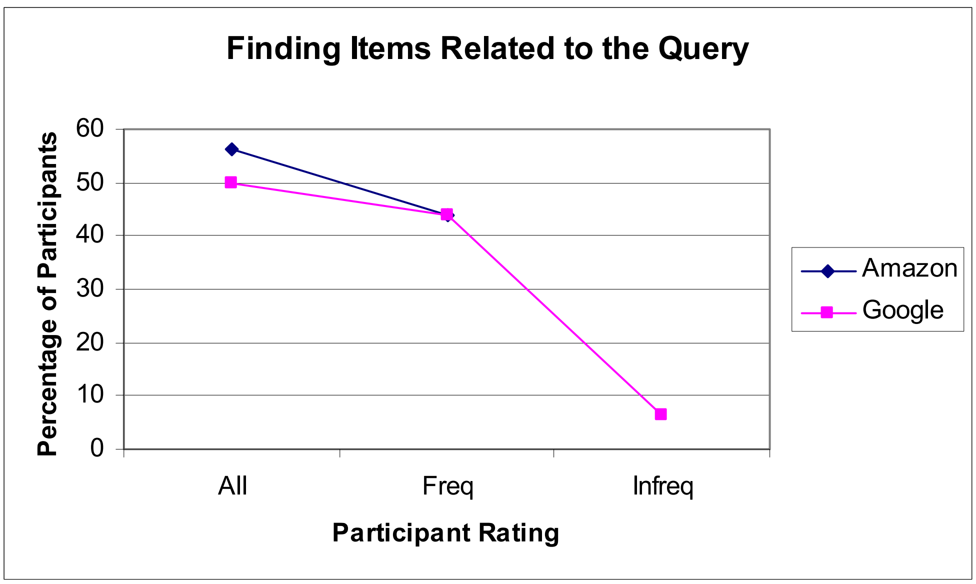

Fig. 6. Comparison of TouchGraph Amazon and Google for Selecting Items

Slight differences between TouchGraph Amazon and Google were shown when participants were asked if they looked at items and decided they were not related to the query. The results favored the TouchGraph Amazon implementation as more participants examined Google-based information and selected the frequently and infrequently categories and fewer participants selected the never category for Google than Amazon (Figure 7).

When participants were asked to rate the Amazon and Google task assignments on the level of difficulty, the results were mixed as more participants rated the Amazon tasks as easy, however more participants rated the Google tasks as moderately easy and of average difficulty (Figure 8). None of the tasks were rated in the difficult or very difficult categories.

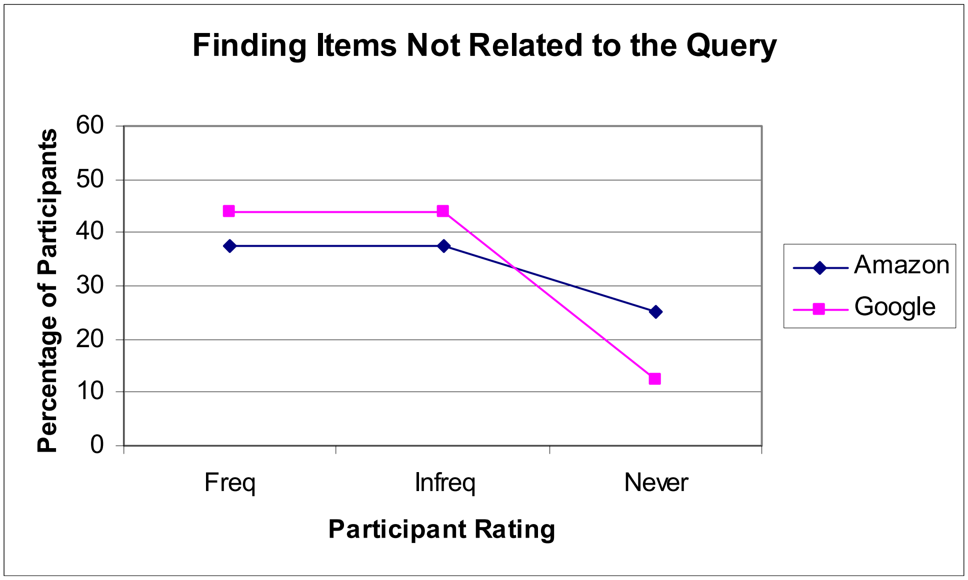

Fig. 7. Comparison of TouchGraph Amazon and Google for Selecting Non-related Items

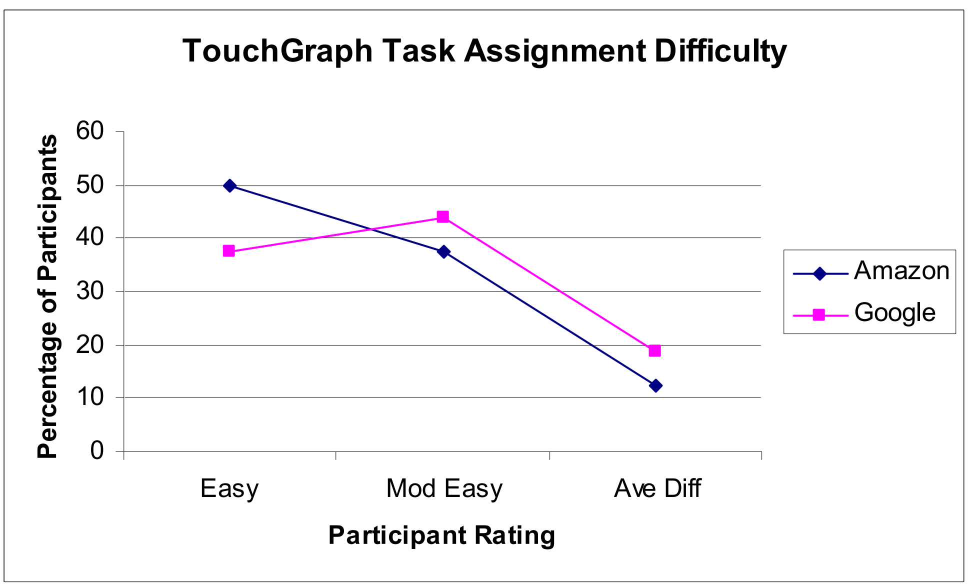

Fig. 8. Participants' Rating of Task Difficulty

Interpreting visualization displays is important to participants' experience with the system and they were asked how they would rate the "understandability" of the TouchGraph display. The highest percentage of participants (47.1%) selected the moderately easy category (Figure 9).

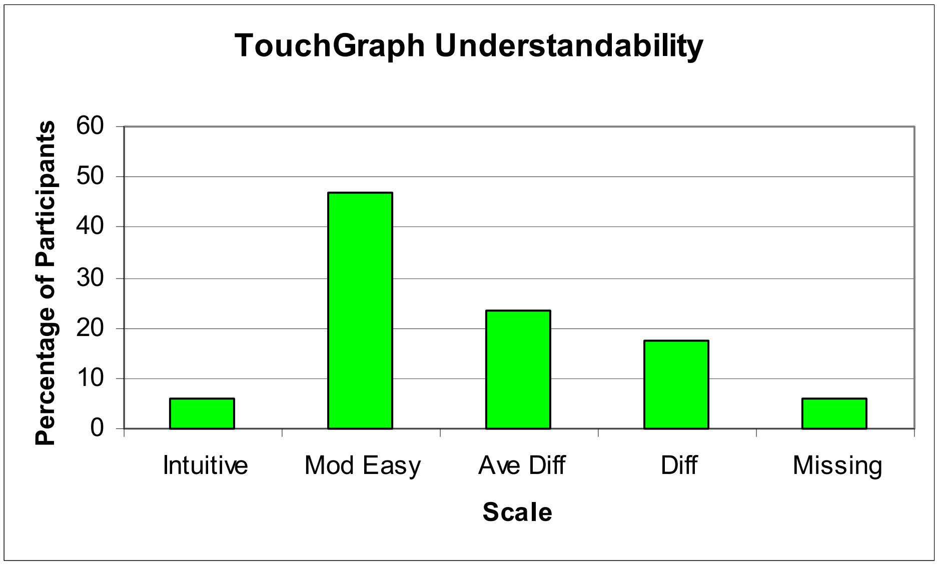

Fig. 9. Participant Rating of the TouchGraph Display Understandability

Finally, the participants were asked what impact they thought that TouchGraph's visual interface has on selecting similar items to a query when compared to a text-based system. The highest percentage of participants (52.9%) selected the somewhat better category (Figure 10)

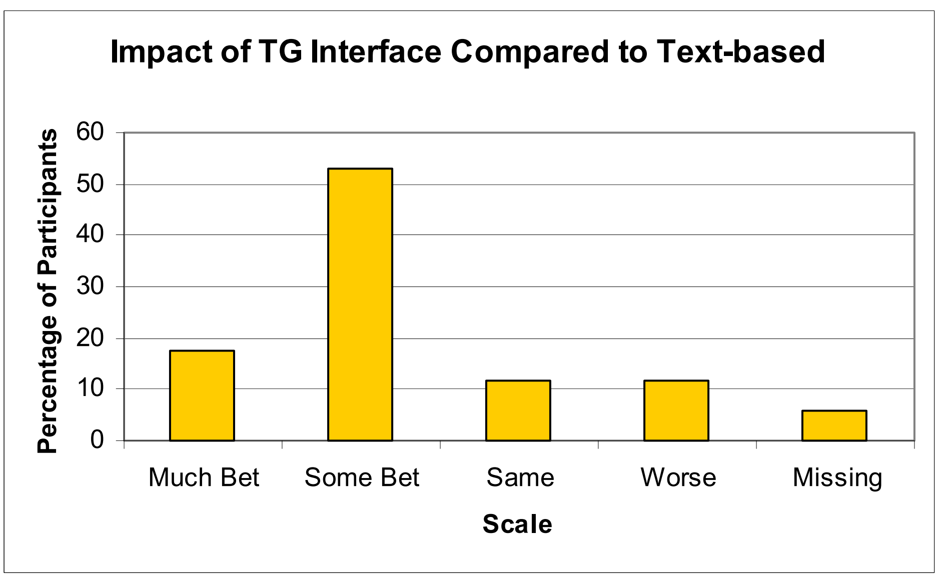

Fig. 10. TouchGraph and Text-based System Comparison

Discussion

The comparison of users selecting search item output based on similarity to the query presented interesting findings. There was a high degree of overlap between the user constructed set and the similarity set generated by the browser, however only a small percentage of participants rated their topic knowledge level as high for two Google tasks. Therefore, these findings do not support the original hypothesis that links greater topic knowledge to a higher overlap between user and system generated sets. Perhaps topic knowledge level is not a significant factor and a visualization display such as TouchGraph has the potential to accommodate a diverse user base for similarity selection.

There was no statistical significance shown between task knowledge level and the use of info buttons or pop-up windows. A statistically significant relationship was found between task knowledge level and node use for three of the six tasks. Part of the original hypothesis that predicted lower use of these features as task knowledge level increased was partially substantiated by these results.

This finding indicated an interesting distinction among the tasks. One possible explanation is that overall these three tasks contained more items identified by users as similar to the query that were not found in the browser-generated set ($C_2 = S - V$). From a visual perspective, these three tasks represented more of a hub and spoke visualization in the initial display; whereas the other tasks presented a more visually dense display. These observations require further analysis.

In regard to subjective data, the participants did not find the tasks difficult, they rated the "understandability" of the display well, and most interestingly, the majority reported that TouchGraph might be better in selecting Web search output than a regular text-based system. While TouchGraph was not compared with a text-based system in this study, this last result is consistent with findings from a prototype visualization-based user study where the visual displays (graph, icon, and "spring"), were favored over textual displays for various tasks [16].

The overlap of user-system item sets and the overall positive evaluations found in the post-search questionnaire raises awareness of the utility of graph visualizations for Web-based retrieval. These representations are referred to as node-link diagrams and they are a familiar concept found in items such as organizational charts and data flows found in software engineering [1]. Common characteristics of this diagram type include uniform sized circles or squares and lines depicting relationships among the entities.

The TouchGraph display may capitalize from the familiarity associated with this particular type of visualization as the relatively homogenous shaped elliptical nodes are understood and the lines are identifiable relationships among the nodes. Essentially, the user does not need to process the meaning of the nodes themselves to interpret the overall visualization. TouchGraph contains options to change the node representation on the display and further research is warranted to test the node diagram (graph) familiarity theory for Web information visualization and interface design.

Limitations

Limitations of the current study include sample size and the use of information retrieval students who may be predisposed to adapting more readily to a new Web-based technology. Although these subjects may be representative of a larger Web user population, future research would attempt to control these factors by increasing the sample size and drawing participants from a more general Web user population. The lack of real queries derived from users is a characteristic limitation of user studies in information retrieval. Tasks were designed to be as realistic as possible, but to provide controlled input in an experimental context.

System Design Implications

From an interface design perspective, the TouchGraph implementation presents a manageable visualization according to the results from objective data for similarity selection and the subjective satisfaction measures. The nodes and "info" buttons were used for most tasks and these represent explicit actions that the participants used as a solution path to the task. Clicking on a node expands it to reveal further links and clicking on the "info" button provided a higher level of detail regarding the item in the node. These interface features were regularly utilized, provided a sufficient level of detail, and were made manifest in the interface design which allowed users to make a similarity selection.

The open-ended responses regarding feature use showed that participants listed items that enabled them to gain perspective on the data set. The zoom and scroll actions were used to navigate the node display. The use of line color indicates the examination of node relationships to determine relationships among items. These findings indicate important considerations in the design of graph-based visualization interfaces in achieving a balance between overview and details as well as between text and visuals.

Conclusions and Further Research

In conclusion, TouchGraph provided a useful visualization design to facilitate similarity selection among users whose levels of topic knowledge varied. TouchGraph offers the potential to improve Web search effectiveness by helping users gain control over the management of Web search results and item selection.

Considerations for future research include evaluating differences in user similarity assessments using TouchGraph in comparison to a text-based search set. Also, comparing user performance with TouchGraph and other graph-based visualizations would enhance the knowledge surrounding user interaction with this type of visual representation on the Web. Web user preferences for selecting similar items with a graph-like visualization such as TouchGraph may be compared to a Web map representation to distinguish the utility of different visualizations for Web-based retrieval.

Overall, the prevalent objective of this research is to understand better the role visualization plays in helping to improve Web search effectiveness and in designing a successful visualization for Web information retrieval. This investigation addresses a compelling new area of testing users with visualization tools for evaluating Web output and lays the foundation for further exploration.

References

1. C. Ware (2004), *Information Visualization: Perception for Design, Second Edition*, San Francisco: Morgan-Kaufmann.
2. E. Rasmussen (2002), *Indexing and retrieval for the Web*, in *Annual Review of Information Science and Technology*, B. Cronin, Editor, Information Today, Inc.: Medford, NJ. pp. 91-124.
3. R. Baeza-Yates and B. Ribeiro-Neto (1999), *Modern Information Retrieval*, New York: ACM Press.
4. T.H. Haveliwala, A. Gionis, D. Klein, and P. Indyk (2002), *Evaluating search strategies for similarity search on the Web*, in *Proceedings of the eleventh international conference on the World Wide Web*. ACM Press: New York. p. 432-442.
5. P. Pirolli, S.K. Card, and M. Van Der Wege (2001), *Visual information foraging in a focus+context visualization*. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Seattle, WA: ACM Press.
6. W. Rivadeneira and B. Bederson (2003), *A study of search result clustering interfaces: comparing textual and zoomable user interfaces*. Retrieved April 15, 2004 from ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/2003-36html/2003-36.htm
7. E. Morse and M. Lewis (1997), *Why information visualization systems sometimes fail*. in *Proceedings of IEEE International Conference on Systems Man and Cybernetics* Orlando, FL.

8. R. Osdin, I. Ounis, and R. White (2002), *Using hierarchical clustering and summarization approaches for Web retrieval: Glasgow at the TREC 2002 Interactive Track*.

9. A. Spoerri (2004), *MetaCrystal: visualizing the degree of overlap between different search engines* in the Proceedings of Thirteenth International *WWW Conference*, New York, NY, May 17-22, pp. 378-379.

10. S. Koshman (2004), *Testing user interaction with a prototype visualization-based information retrieval system.* Journal of Documentation, Vol 60, pp. 265-280.

11. A. Veerasamy and N.J. Belkin (1996), *Evaluation of a tool for visualizaton of information retrieval results*. in *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Zurich: ACM Press, pp. 85-92.

12. S.K. Card, J.D. Mackinley, and B. Shneiderman (1999), *Focus+Context*, in *Readings in information visualization: using vision to think*, S.K. Card, Mackinley, J.D. & Shneiderman, B., Editors. Morgan Kaufmann Publishers: San Francisco. p. 307-309.

13. P. Kavassalis, S. Lelis, M. Rafe, and S. Haridi (2004), *What makes a Web site popular.* Communications of the ACM, Vol. 47, pp. 51-55.

14. B. Shneiderman and C. Plaisant (2004), *Designing the User Interface: Strategies for Effective Human-Computer Interaction, Fourth Edition*, Boston: Pearson Addison Wesley.

15. A. Spink, D. Wolfram, B.J. Jansen, and T. Saracevic (2001), *Searching the web: the public and their queries.* Journal of the American Society for Information Science and Technology, Vol 52, pp. 226-234.

16. E. Morse, M. Lewis, and K. Olsen (2002), *Testing visual information retrieval methodologies case study: comparative analysis of textual, icon, graphical and "spring" displays.* Journal of the American Society for Information Science and Technology, Vol 53, pp. 28-40.