



## Representing User Navigation in XML Retrieval with Structural Summaries

Ali, M. S.; Consens, Mariano P.; Larsen, Birger

DOI:  
[10.1007/978-3-642-00958-7\\_73](https://doi.org/10.1007/978-3-642-00958-7_73)

Publication date:  
2009

Document version  
Early version, also known as pre-print

Citation for published version (APA):  
Ali, M. S., Consens, M. P., & Larsen, B. (2009). *Representing User Navigation in XML Retrieval with Structural Summaries*. Paper presented at European Conference on IR Research, Toulouse, France.  
[https://doi.org/10.1007/978-3-642-00958-7\\_73](https://doi.org/10.1007/978-3-642-00958-7_73)

# Representing User Navigation in XML Retrieval with Structural Summaries

M. S. Ali<sup>1</sup>, Mariano P. Consens<sup>1</sup>, and Birger Larsen<sup>2</sup>

<sup>1</sup> University of Toronto, Canada

<sup>2</sup> Royal School of Library and Information Science, Denmark  
{sali, consens}@mie.utoronto.ca, blar@db.dk

**Abstract.** This poster presents a novel way to represent user navigation in XML retrieval using collection statistics from XML summaries. Currently, developing user navigation models in XML retrieval is costly and the models are specific to collected user assessments. We address this problem by proposing summary navigation models which describe user navigation in terms of XML summaries. We develop our proposal using assessments collected in the interactive track at INEX 2006. Our preliminary results suggest that summary navigation models can represent user navigation in a way that is effective for evaluation and allows economic re-use of assessments for new tasks and collections.

## 1 Introduction

XML retrieval systems focus results by returning XML elements or text passages that help users navigate to relevant information in retrieved documents. User navigation models represent how users navigate to relevant information in documents. They are based on assessments collected in costly user studies. A significant problem is that these assessments are specific to the user study and often not re-usable for new topics, tasks or collections. To our knowledge, there are no existing approaches to alleviate this problem. Our proposal is to use summary navigation models which model user navigation in a way that can be calculated without assessments, and is amenable to the evaluation of new tasks, topics and collections. Section 2 summarizes our observations from the interactive track user study at INEX 2006. Section 3 presents user navigation models based on the user study. In Section 4, we propose a number of different summary navigation models. Section 5 concludes by comparing how well our proposed summary models agree with user models from INEX 2006.

## 2 Interactive Track at INEX 2006

The interactive track at INEX<sup>3</sup> is an international, multi-year user study to identify XML elements that are helpful in solving given search tasks. The 2006 user study consisted of 83 participants for 12 assessment topics with user activity recorded for 818 documents from the INEX 2006 Wikipedia collection [3].

---

<sup>3</sup> INitiative for the Evaluation of XML retrieval (INEX): <http://www.inex.otago.ac.nz>

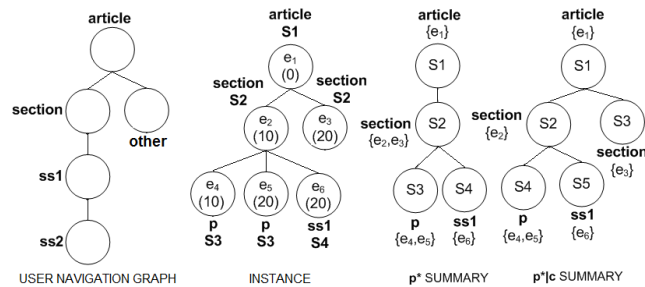


Fig. 1. User navigation graph (first) and Summary example (second to fourth).

Source	Destination				
	ARTICLE	SEC	SS1	SS2	OTHER
ARTICLE	0 (0)	138 (100.4)	18 (48.7)	1 (22)	2 (76)
SEC	278 (57.0)	372 (14.7)	41 (11.3)	0 (0)	0 (0)
SS1	46 (13.1)	50 (10.2)	50 (9.52)	0 (0)	1 (48)
SS2	4 (12.3)	2 (264.5)	13 (5.3)	0 (0)	0 (0)
OTHER	7 (27.7)	0 (0)	1 (4)	0 (0)	4 (26)

Table 1. Number of visits (mean time spent)

The interactive track prototype XML retrieval system is a scrollable window that contains a frame to display the entire article (with the returned elements highlighted in context), and a separate frame that displays the table-of-contents as a set of links. User events are time-stamped and the duration of an event is the length of time between its start time and the start time of the next event in the participant's session. To view articles for assessment, the participant clicks on a result from a list of hyperlinks (DETAIL QUERY). The system does not track navigation via the scrolling behavior of users. It does however track whether the participant uses the table-of-contents to navigate to different parts of the article (DETAIL BROWSE). A visit to an element is defined as a participant who enters an element via a DETAIL QUERY or a DETAIL BROWSE.

The first graph in Figure 1 shows the five types of XML elements that participants visited in the 2006 user study; namely, ARTICLE, SEC, SS1, SS2, and OTHER. These correspond to elements whose label paths are the root */article* (ARTICLE), a section path */article/body/section* (SEC), a subsection path *SEC/section* (SS1), a sub-subsection path *SS1/section* (SS2), and all other elements' paths (OTHER). We call Figure 1(first) the user navigation graph. Table 1 tabulates the visits and mean time spent in visits for element assessments. For instance, participants visited SS2 elements and then navigated to element ARTICLE 4 times. The mean time spent in SS2 before navigating to element ARTICLE was on average 12.3 seconds. This led to an overall time, which we refer to as an episode, of  $12.3 \times 4 = 49.2$  seconds. The most visited element was SEC, and the largest mean time spent occurred in navigations to SEC elements from ARTICLE.

### 3 User Navigation Models

User navigation models represent one aspect of the information seeking behaviour of users. In this work, our model is the probability  $\pi_{(e)}$  that the user will navigate to element  $e$  if they enter the document that contains element  $e$ . We determine user navigation, given a user navigation graph, by first weighting the paths in the graph, then normalizing these weights in a matrix, and finally finding  $\pi_{(e)}$  by iteratively multiplying the matrix with itself, until all rows are equal (*i.e.*, steady-state probabilities [5]).

Consider row **SS2** in Table 1. The total number of navigations from **SS2** elements is the sum of the row,  $4 + 13 + 2 + 0 + 0 = 19$ . We normalize the row by dividing the entire row by its sum, so the normalized row **SS2** is  $4/19 = 0.21$  for **ARTICLE**,  $2/19 = 0.11$  for **SEC**,  $13/19 = 0.68$  for **SS1**, and  $0/19 = 0.0$  for both **SS2** and **OTHER**. Table 2A shows the resultant normalized weighting matrix for all rows based on the number of visits. Now, to get a user navigation model, we iteratively multiply the normalized weighting matrix with itself until all rows are equal. Table 2B shows models based on the number of visits (**Visit**), the overall amount of time spent in elements (**Episode**), and the mean time spent in each element (**Time spent**). In the next section we derive similar models but instead weight the user navigation graph with XML structural statistics.

#### A. Normalized Weights for Visits

Source	Destination				
	ARTICLE	SEC	SS1	SS2	OTHER
ARTICLE	0.0	0.87	0.11	0.01	0.01
SEC	0.40	0.54	0.06	0.0	0.0
SS1	0.31	0.34	0.0	0.0	0.01
SS2	0.21	0.11	0.68	0.0	0.0
OTHER	0.58	0.0	0.08	0.0	0.33

#### B. User Navigation Models

	ARTICLE	SEC	SS1	SS2	OTHER
Visit	0.281	0.606	0.105	0.002	0.006
Episode	0.410	0.531	0.050	0.001	0.009
Time spent	0.318	0.209	0.129	0.028	0.317

#### C. Summary Navigation Models

	ARTICLE	SEC	SS1	SS2	OTHER
Path	0.361	0.537	0.087	0.014	0.001
Content	0.103	0.434	0.089	0.013	0.361
Depth	0.309	0.435	0.067	0.008	0.181

**Table 2.** Visit user model transition matrix.

## 4 Summary Navigation Models

XML structural summaries are graphs representing relationships between sets of document elements with a common structure (paths, subtrees, etc.). For instance, AxPRE summaries [2] define a broad range of the different summaries available in the literature. They are created using an axis path regular expression language that is capable of describing a plethora of partitioning schemes. For example, a  $p^*$  summary partitions XML elements based on their incoming paths, since  $p^*$  is the axis path regular expression describing paths of parent ( $p$ ) axis traversals. Similarly, a  $p^*|c$  summary is the axis path regular expression describing paths of parent ( $p$ ) with a single child ( $c$ ) axis traversals. Figure 1 shows an example Wikipedia article instance (second), its  $p^*$  summary (third), and its  $p^*|c$  summary (fourth). The elements in the summary partitions are called the *extent* of the summary partition.

Summaries provide a formal way to represent user navigation graphs based on the structure of documents in the collection. Summary navigation models ascribe weights to the edges (as opposed to paths) of the user navigation graph that are derived from the structural properties (such as the content length or label path depth) of child nodes. We consider three different weighting schemes. Path weights are the extent size of the child nodes in the summary of the user navigation graph. Content weights are the number of characters of content in the elements in the extent of the child nodes. Finally, depth weights are the same as content but damped (divided) by the path depth of the elements in the extent of the child nodes. Using the methodology in the previous section and 2343 randomly selected Wikipedia articles summarized using a  $p^*$  summary which was then mapped to the user navigation graph shown in the first graph in Figure 1, Table 2C shows the resulting summary navigation models based on path, content and depth weights.

## 5 Results

The user and summary navigation models shown in Table 2B and Table 2C were used to evaluate Wikipedia runs using mean-average Structural Relevance in Precision. Structural relevance (SR) is the expected relevance of a ranked list given that retrieved elements are redundant to the user [1]. An element is redundant (and thus non-relevant) if the user sees it more than once from the list. It is calculated by conditioning the relevance of each element in the list with the probability that the user will navigate to it.

$$SR(R) = \sum_{i=1}^k rel(e_i) \cdot \pi_{(e_i)}^{m(R,e_i)} \quad (1)$$

where  $R$  is a ranked list of  $k$  elements,  $e_i$  is the  $i$ -th element in the list,  $rel(e)$  is the relevance value of the element  $e$ ,  $\pi_{(e)}$  is the user navigation probability that element  $e$  will be navigated to if the user enters the document that contains

$e$ , and  $m(R, e)$  is the number of higher-ranked elements from the same document as  $e$ . We evaluate systems using Structural Relevance in Precision (SRP) which is  $SR(R)/k^4$ .

The experimental setup was the INEX 2006 Wikipedia collection for 15 systems across 107 topics in the Ad-hoc Thorough task [4] at rank cut-offs of  $k=10$  and  $k=50$ . The systems were ranked across topics using SRP parameterized with the 6 different navigation models (visit, episode, time spent, path, content and depth). The 6 system rankings were then compared using Spearman’s Rho p-value correlations (p-value < 0.1 meant correlated rankings, and p-value < 0.05 meant strongly correlated rankings). Table 3 shows the p-value (correlations) between user navigation models and our proposed summary navigation models.<sup>5</sup>

k=10 (k=50) <i>Summary Models</i>	<i>User Models</i>		
	<b>path</b>	<b>content</b>	<b>depth</b>
<b>episode</b>	0.005 (0.010)	0.099 (0.127)	0.037 (0.061)
<b>visit</b>	0.004 (0.012)	0.111 (0.118)	0.054 (0.039)
<b>time spent</b>	0.109 (0.087)	0.033 (0.062)	0.043 (0.058)

**Table 3.** Correlation (p-value) of User and Summary Nav. Models for  $k=10$  ( $k=50$ )

The path model had strong correlation (p-value < 0.05) with both the episode and visit user models. The content model showed correlation (p-value < 0.1) with the time spent model. Overall, the depth model demonstrated the best results, in that, it showed correlation with all user models for both  $k = 10$  and  $k = 50$ . These results suggest that the depth model could be used as a general user navigation model. The strength of our approach is that summary navigation models can be economically applied to new collections, because they are only validated using assessments, but calculated using summaries and structural properties of the elements in the collection.

## References

1. M. S. Ali, M. P. Consens, G. Kazai, and M. Lalmas. Structural Relevance: A common basis for the evaluation of structured document retrieval. In *CIKM 2008*, pages 1153–1162, 2008.
2. M. P. Consens, F. Rizzolo, and A. A. Vaisman. AxPRE Summaries: Exploring the (Semi-)Structure of XML Web Collections. In *ICDE 2008*, pages 1519–1521, 2008.
3. S. Malik, A. Tombros, and B. Larsen. The interactive track at INEX2006. *LNCS 4518*, pages 387–399, 2007.
4. S. Malik, A. Trotman, M. Lalmas, and N. Fuhr. Overview of INEX 2006. *LNCS 4518*, pages 1–11, 2007.
5. S. M. Ross. *Introduction to Probability Models*. Academic Press, 2003.

<sup>4</sup> In [1], SRP was determined to be accurate for XML retrieval evaluation using the INEX 2006 Wikipedia collection in the Ad-hoc Thorough Task.

<sup>5</sup> Spearman’s Rho was greater than 0.5 for all comparisons.