Explaining Search Results*

Maurice Coyle and Barry Smyth
Smart Media Institute, Department of Computer Science,
University College Dublin, Ireland
{maurice.coyle, barry.smyth}@ucd.ie

Abstract

In this paper we argue that it may be possible to help searchers to better understand the relevance of search results by generating explanations that highlight how other users have interacted with such results under similar search conditions in the past. We propose the use of the search histories of a community of online users as a source of these explanations. We describe the results of a recent study to examine the use of such explanation-based techniques to help Web searchers better appreciate the relevancy of search results. We highlight shortcomings of this approach in its current form and offer suggestions as to how it may be improved in future work.

1 Introduction

Scale, document diversity, and limited searcher expertise all combine to make the Web a very challenging information retrieval environment. Consequently, Web search engines have had to evolve beyond their term-based information retrieval roots; see [Brin and Page, 1998; Lawrence, 2000] for approaches that leverage connectivity and context information to improve search. In this paper we adopt a very different perspective on how search engines might be improved. We ask whether there is an opportunity to present search results to users more effectively. This is related to the use of clustering methods to impose order on search results at presentation time (e.g., [Leuski, 2001]). However, we try to *explain* results to users, to help users gain a better understanding of the appropriateness of these results to their needs.

In related work, [Hearst, 1995] introduces a visualisation paradigm that offers an explanation not just of how strong the match between the user's query and a given result page is, but also how frequent each term is, how each term is distributed in the text and where the terms overlap within the document. This approach provides the user with additional explanatory information which better aids them in their decision as to whether or not each page is relevant to their information need. For related work on the issue of explanation in recommender systems see [Cunningham and McSherry, 2004].

Of course, Web search engines do attempt to provide *some* explanation with their result-lists. Typically, for example, each result is accompanied by a *snippet* of text from the result page in question to give some insight into the content of the page. Both these query-sensitive snippets and the approach to explanation in information retrieval introduced in the TileBars system ([Hearst, 1995]) are content-based techniques for aiding the user in result selection. We believe that other forms of explanatory information can be usefully provided.

In [Wexelblat and Maes, 1999], *interaction histories* are introduced, whereby a user's interactions with their Web browser (e.g. selection of a link in a Web page, typing a URL into the location bar, etc.) are recorded as they browse, and this information is later used to help other users to locate information items. The concept of interaction histories is applied to Web search in [Coyle and Smyth, 2005]; the pages in a user's navigation sequence beyond the initial results page are recorded and used to enhance future search result lists when similar queries are submitted. We propose the use of search history information (specifically query-submission and result -selection information) for the generation of search result explanations. In this paper we describe a recent study to evaluate different approaches to generating these explanations based on 3 types of user interaction information.

2 Explaining Search Results

Result snippets provide one type of explanation. What other types of explanation might be generated? In this paper we consider 3 additional types of explanation information that can be derived from the search histories of communities of users. We will describe these types of information in a moment but first it is worthwhile considering the source of these search histories.

2.1 Collaborative Web Search

Collaborative Web search (CWS) is a form of meta-search, relying on the search services of a set of underlying search engines and manipulating their results in response to the learned preferences of a given community of users ([Smyth *et al.*, 2003; In Press]). Very briefly, a central data structure in CWS is the *hit-matrix*, H. It encodes the search behaviour of a given community of users; each time a community member

*The support of the Informatics Research Initiative of Enterprise Ireland is gratefully acknowledged.

selects a result p_j in response to some query q_i, the entry in cell H_{ij} is incremented. In turn, the *relevance* of a page p_j to q_i can be estimated as the relative number of selections p_j has received in the past for q_i. For example, if the page *www.jaguarcars.com* has a relevance of 0.65 for the query *jaguar*, this means that 65% of the selections for this query have been for this result. More generally, the relevance of a page p_j to some target query q_T can be calculated as the weighted sum of its relevance to a set of queries $(q_1, ..., q_n)$ that are similar to q_T, with each individual relevance value discounted by the similarity between q_T and the query in question; see [Smyth *et al.*, In Press] for further information.

The type of interaction data generated by CWS, which can be stored and later used as a source of search histories includes, but is not limited to, search result selection, query submission and query modification (i.e., the same user submitting a number of queries one after the other which are similar to each other regarding their component terms). The point to note is that this stored interaction data can be leveraged in future search sessions to provide explanation text alongside search results which can be used to determine a result's popularity, by analysing how users within the same searching community have previously interacted with it. With this added information, the user is better positioned to judge whether or not each result is relevant to their information need.

2.2 Explanation Types

In this paper we propose to use the interaction data captured by CWS as a source of 3 different explanation types: 1) relevance scores; 2) related queries; 3) timing information. From these we have generated the following 6 explanation types.

Type 1 - Relevance. The relevance of a result to some query is simply the percentage of times that other searchers have selected the page for the query. Thus each result is accompanied by an explanation of the form: *"r% of searchers for this query have selected this page"*.

Type 2 - Related Queries. Related queries refer to other queries that have also led to the selection of the page in question. They provide users with an alternative understanding of the circumstances in which other users have found a page to be worth selecting; e.g., *"This page has also been selected for queries such as 'q1' and 'q2'"*.

Type 3 - Timing. Timing information helps the searcher to understand the recency of any selection activity for this page and may give an indication of the currency of its content; e.g., *"This page was last selected t minutes ago"*.

Type 4 - Relevance + Related Queries. We can combine relevance information with related queries to generate explanations of the form: *"r1% of searchers for this query have selected this page and r2% of searchers for similar queries such as 'q1' and 'q2' have also selected it"*.

Type 5 - Relevance + Timing. Similarly, we can combine relevance and timing information to generate explanations of the form: *"r% of searchers for this query have selected this page as recently as t minutes ago"*.

Type 6 - Relevance + Related Queries + Timing. And finally, we can combine all 3 types of data to produce explanations of the form: *"As recently as t minutes ago, r1% of searchers for this query have selected this page and r2% of searchers for similar queries such as 'q1' and 'q2' have also selected it"*.

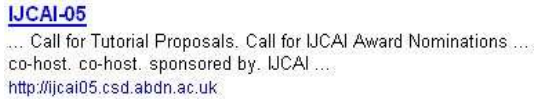

Figure 1: Search result for the query 'ijcai' with only a snippet text explanation.

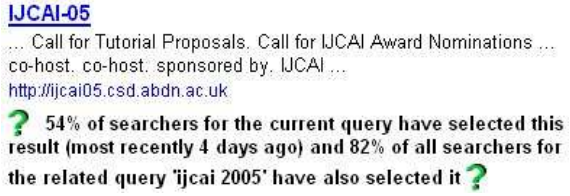

Figure 2: Search result for the query 'ijcai' with interaction history-based explanation (type 6) added below the snippet text explanation.

By way of an example, consider the search results presented in Figures 1 and 2. Figure 1 shows a standard Web search result returned for the query 'ijcai' with the page's title and a contextualised snippet provided as a means for the user to determine if the result is relevant to their query. In Figure 2, we can see the same search result but this time with explanation text added below it (explanation type 6, see above) which was generated using the interaction histories of previous users. This explanation contains information about how many times this result has been selected in the past for the current query and also for other, related queries. Further, the user can see how recently this result has been selected, which should give them some idea of how 'fresh' the result is for the current query.

3 Evaluating Search Result Explanations

To evaluate the utility of these different explanation types, we surveyed the opinions of 57 postgraduate students. Each survey was composed of 6 pairs of search engine result-lists. In each pair there was a *standard* result-list, that contained the sort of result-list presented by Google and other major search engines; each result contained a title, a snippet, and a URL (see Figure 1). Each pair also contained an *explanation* result-list in which the *same* results were presented with one of the 6 explanation types described above (see Figure 2). The 6 pairs of each survey allowed users to compare each explanation type to a standard result-list; for each pair the user was asked to indicate whether they felt the explanations helped them to understand the value of the results in relation to the target query. In each survey the type of results and queries used to generate them were randomised so no two pairs of result-lists contained the same query-result combination for any user.

Summary results are presented in Figure 3 as the percentage of users who found the explanations to be *more informative* (better) than the standard result-list against each result type; we also show the percentage of users who found each

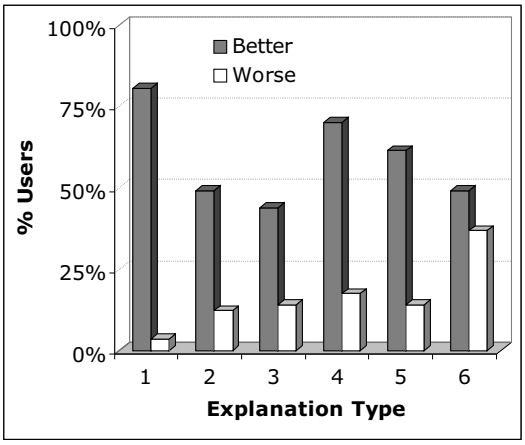

Figure 3: Summary of results.

explanation type to be *less informative* (worse) than the default. The results show a clear preference for relevance information (type 1 explanations) with more than 80% of users indicating that this type of explanation helped them to better understand the usefulness of results; only 3% of users rated this explanation type as worse than the default.

In general, the users appear to be less convinced about the value of the other explanation types with just under 50% viewing the related queries (type 2) and timing (type 3) as better than the default; although only about 12%-14% of users found these explanations to be worse than the default. When these explanation types are combined with relevance information (in types 4, 5 and 6) the user ratings are seen to increase, although never above the 80% found for relevance information on its own. This increase appears to be due to the inclusion of the relevance data, however.

4 Conclusions

In this paper we have suggested that modern Web search engines might be usefully improved if they included some form of explanation information alongside their search results. This could potentially provide searchers with a clearer understanding of how each result relates to their information need, if at all. We have tested a number of different explanation types using a combination of information sources.

The results of this experiment indicate that there is potential for Web searchers to benefit from the presentation of additional explanatory information alongside standard result-list information. Concrete relevance information in particular was found to be valued by more than 80% of users, although other forms of explanation were found to be less compelling. One of the problems with these other forms of explanation data appears to be linked to their increased space requirements. For example, the use of related queries and timing information in particular adds considerably to the size of explanations, which takes away from valuable screen-space that could otherwise be used to present additional search results.

It is likely that this presentation issue could be resolved however, either through the use of more economical textual presentation techniques such as mouse-over tool-tips or else through the use of a visualisation approach whereby the user interaction information for a search result is summarised using graphical icons in a manner similar to that found in [Hearst, 1995]. It is worth noting that only a small percentage of users viewed the additional explanations as having a negative impact on the search results. This allows us to be optimistic that a more economical, visual presentation metaphor might reduce this percentage and lead to an increase in the perceived worth of the various explanation types.

References

[Brin and Page, 1998] Sergey Brin and Lawrence Page. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30(1–7):107–117, 1998.

[Coyle and Smyth, 2005] Maurice Coyle and Barry Smyth. Enhancing web search result lists using interaction histories. In *Proceedings of the 27th European Conference on Information Retrieval, (ECIR-05)*, pages 543–545, Santiago De Compostela, Spain, 2005. Springer-Verlag.

[Cunningham and McSherry, 2004] Padraig Cunningham and David McSherry. *Workshop on Explanation in CBR*. 7th European Conference on Case-Based Reasoning, Madrid, Spain, 2004.

[Hearst, 1995] Marti A. Hearst. Tilebars: Visualisation of term distribution information in full text information access. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'95*, pages 59–66, Denver, Colorado, USA, May 1995. ACM Press.

[Lawrence, 2000] Steve Lawrence. Context in Web Search. *IEEE Data Engineering Bulletin*, 23(3):25–32, 2000.

[Leuski, 2001] Anton Leuski. Evaluating document clustering for interactive information retrieval. In Ling Liu Henrique Paques and David Grossman, editors, *Proceedings of Tenth International Conference on Information and Knowledge Management (CIKM'01)*, pages 41–48, Atlanta, Georgia, USA, 2001. ACM Press.

[Smyth *et al.*, 2003] Barry Smyth, Evelyn Balfe, Peter Briggs, Maurice Coyle, and Jill Freyne. Collaborative Web Search. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, (IJCAI-03)*, pages 1417–1419. Morgan Kaufmann, 2003. Acapulco, Mexico.

[Smyth *et al.*, In Press] Barry Smyth, Evelyn Balfe, Jill Freyne, Peter Briggs, Maurice Coyle, and Oisin Boydell. Exploiting Query Repetition and Regularity in an Adaptive Community-based Web Search Engine. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, (In Press).

[Wexelblat and Maes, 1999] Alan Wexelblat and Pattie Maes. Footprints: history-rich tools for information foraging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'99*, pages 270–277, Pittsburgh, Pennsylvania, USA, 1999. ACM Press.