Proceedings of the 3rd Workshop on Intelligent Techniques for Web Personalization (ITWP'05), in conjunction with the 19th International Joint Conference on Artificial Intelligence (IJCAI'05), Edinburgh, Scotland, UK. August 1, 2005. (Bamshad Mobasher and Sarabjot Singh Anand, editors).

Communities, Collaboration and Cooperation in Personalized Web Search*

Jill Freyne and Barry Smyth
Department of Computer Science,
University College Dublin,
Ireland.
{firstname.lastname}@ucd.ie

Abstract

Collaborative Web Search is an approach to Web search that takes advantage of ideas from case-based reasoning in order to improve upon the results of a standard search engine. It retrieves and reuses the prior search experiences of a community of like-minded searchers. Collaborative Web search contemplates a *society* of differently focused communities, each with their own case-base of search experiences. In the past we have focused on how the experiences of a single community can assist with future searches by its community's members, by personalizing results according to learned community preferences. This paper borrows ideas from collaborative and distributed case-based reasoning to show how search experiences of related communities can also usefully assist search.

1 Introduction

Web search is challenging for many reasons, not least of which is the sheer scale and heterogeneity of the Internet. Moreover, the average searcher is far from the information retrieval expert assumed by the information retrieval techniques underpinning today's search engines. For example, most Web search queries are notoriously vague and contain 2-3 terms [Lawrence and Giles, 1998] that are often poorly chosen with respect to the documents being sought [Furnas *et al.*, 1987].

Consequently, the field of Web search has seen many recent developments in order to adapt traditional term-matching information retrieval techniques for the needs of today's Web searchers. This has led to a number of significant innovations, which are now part of mainstream search. For example, the works of [Brin and Page, 1998] and [Kleinberg, 1999] have famously exploited document connectivity information to significantly improve retrieval performance. More recently, and in specific response to the vague-query problem mentioned above, other researchers have sought to exploit context as a means of supplementing vague queries and so guiding search better [Lawrence, 2000]. Alternatively, other researchers have looked at clustering techniques as a way to impose order on a collection of search results, with a view to identifying different conceptual groupings of results [Dell Zhang, 2004; Hamilton, 2003].

In our work we adopt a very different approach to improving the quality of Web search. By borrowing ideas from case-based reasoning [Aamodt and Plaza, 1994] and social networking [Scott, 1991], we have developed *Collaborative Web Search*, which operates as a form of personalized meta-search, exploiting the resources of a traditional search engine. In addition, the search histories of a community of like-minded searchers are recorded as *search cases* in a case-base. These communities can be loosely or more formally defined. For example, a community might be all of those searchers that use a search box on a motoring Web site, their search queries stored in an associated case-base. Each case corresponds to a specific past query and the resulting selections. These cases are reused in response to similar target queries, with past result selections promoted to the searcher. Previously, we have shown how this approach can significantly improve the precision and recall characteristics of existing Web search engines [Smyth *et al.*, 2004].

Collaborative Web search contemplates a society of community-based search engines, each with their own case-base of search cases corresponding to a distinct community of searchers. Ordinarily the searches of a specific (*host*) community are answered with reference to their local case-base; so for searches from our motoring Web site, promotions are made on the basis of previous similar queries and selected results that have come through this specific community. However, in this work we take this idea a step further, drawing on ideas from collaborative and distributed case-based reasoning. In particular, we consider the possibility of leveraging the search experience of other related communities when responding to the queries of a host community, on the basis that these related communities may offer a useful alternative perspective for our searcher. Consider the example above of a general motoring Web site. A visitor might submit the query *'jaguar'* and the host community might promote pages to do with Jaguar cars because these results have been previously selected for similar queries. However, there may also be a more specialised community based around *Jaguar-Enthusiasts.com* whose search case-base may have other relevant results, perhaps results that emphasise restoration work for example, to recommend to our target searcher.

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

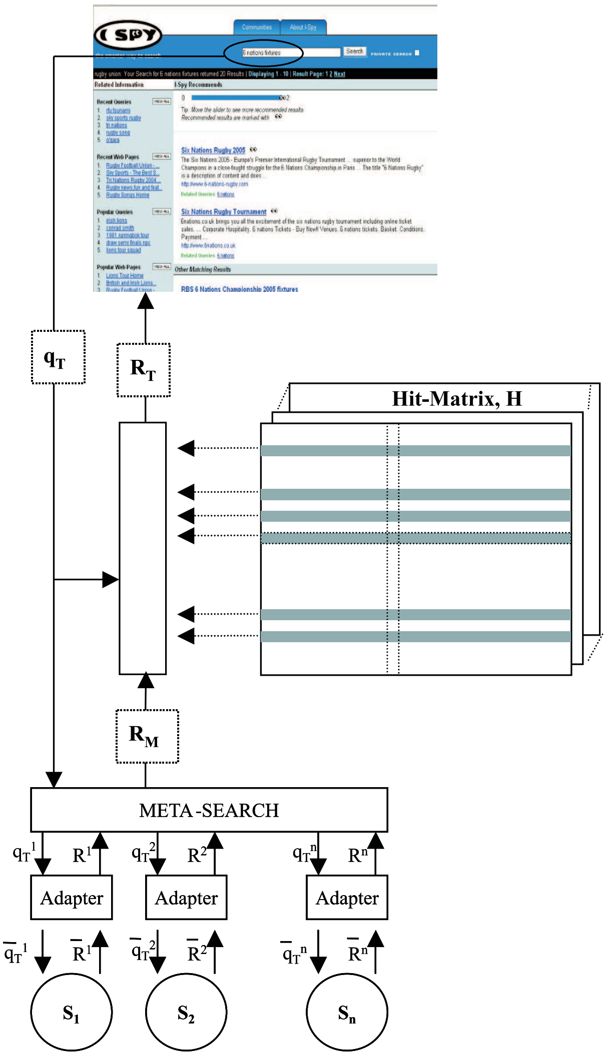

Figure 1: Collaborative Web search as implemented in I-SPY.

In this paper, after providing a review of collaborative Web search, we describe (in Section 3) how it can be adapted to take advantage of community overlaps with a view to improving the results that are ultimately recommended to the user. In particular, we describe how related communities can be identified and how their results can be selected and scored relative to the host community. Before concluding with a discussion of related work, Section 4 describes the results of an experiment that helps to clarify the relative advantages of this approach within a particular cluster of related communities.

2 A Review of Collaborative Web Search

Collaborative Web search is a form of meta-search; see Figure 1. Each new query, q_T, is submitted to a set of underlying search engines and their results are combined to form a meta-search result-list, R_M. The key novelty stems from how this result-list is processed to produce a new result-list, R_T, that reflects the learned preferences of a community of like-minded searchers. This involves reusing selection results from past search cases for similar queries, promoting those results that were reliably selected in the past.

2.1 Profiling Community Preferences

The *hit-matrix*, H, is a key data structure for collaborative Web search, which relates page selections to past queries for a community of users. Specifically, H_{ij} (the *hit value* of page p_j for query q_i) is the number of times that page p_j has been selected for query q_i; H_{ij} is incremented each time p_j is selected for q_i. The hit-matrix forms the basis of a case-base. Of course we do not implement this structure as a matrix but rather use a set of Oracle tables to implement a more efficient and dynamic record structure that is better suited to our needs. Nevertheless, its suits us to consider this structure as a matrix for the purpose of explanation. Each row corresponds to a search case (see Equation 1) or, equivalently, a $k+1$-tuple made up of the query component (a set of query terms) plus k result-pairs, each with a page id and an associated percentage relevance value computed from the hit value for this page and query combination; we explain how this relevance value is computed in Equation 6. The *problem specification* part of the case (see Equation 2) corresponds to the query terms. The *solution* part of the case (see Equation 3) corresponds to the result-pairs; that is, the set of page selections recorded as a result of past uses of the corresponding query. The *target problem* is, of course, represented by the target query terms.

$$c_i = (q_i, (p_1, r_1), ..., (p_k, r_k)) \tag{1}$$

$$Spec(c_i) = q_i \tag{2}$$

$$Sol(c_i) = ((p_1, r_1), ..., (p_k, r_k)) \tag{3}$$

$$Rel(p_j, c_i) = r_j \texttt{ if } (p_j, r_j) \in Sol(c_i); = 0, otherwise. \tag{4}$$

2.2 Retrieving Similar Search Cases

For each target query q_T we retrieve a set of similar search cases to serve as a source of relevant results. Case similarity can be measured using a simple term-overlap metric (Equation 5); alternative similarity metrics have been recently evaluated [Balfe and Smyth, 2005] but for now we will use this simple measure. During the retrieval stage, this allows collaborative search to rank-order past search cases according to their similarity to the target query so that all, or a subset of, these similar cases might be reused during result ranking.

$$Sim(q_T, c_i) = \frac{|q_T \cap Spec(c_i)|}{|q_T \cup Spec(c_i)|} \tag{5}$$

2.3 Case Reuse & Result Ranking

Consider a page, p_j, that is part of the solution of a case, c_i, with query, q_i. The relevance of p_j to this case is given by the relative number of times that p_j has been selected for q_i; see Equation 6. And the relevance of p_j to the current target query, q_T, is the combination of $Relevance(p_j, q_i)$'s for all pages that are part of the solutions to cases $(c_1, ..., c_n)$ deemed to be similar to q_T, as shown in Equation 7. Essentially each $Relevance(p_j, q_i)$ is weighted by $Sim(q_T, c_i)$ to discount the relevance of results from less similar queries; $Exists(p_j, c_i) = 1$ if $H_{ij} <> 0$ and 0 otherwise.

$$Relevance(p_j, q_i) = \frac{H_{ij}}{\sum_{\forall j} H_{ij}} \tag{6}$$

$$WRel(p_j, q_T, c_1, ..., c_n) = \frac{\sum_{i=1,...,n} Rel(p_j, c_i) \bullet Sim(q_T, c_i))}{\sum_{i=1,...,n} Exists(p_j, c_i) \bullet Sim(q_T, c_i)} \tag{7}$$

This weighted relevance metric is used to rank promotion candidates. These ranked pages are then recommended ahead of the remaining meta-search results, which are themselves ranked (according to a standard meta-search scoring metric), to give R_T. Alternative promotion models can also be envisaged but are omitted here due to space constraints.

3 Communities, Collaboration and Cooperation

Obviously the success of collaborative Web search depends on the extent to which the contents of a given hit-matrix (search case-base) reflect some relatively uniform domain of interests, for a community of like-minded searchers. In reality this is relatively commonplace since many search scenarios correspond to a form of community search, and within these scenarios there tends to be a strong degree of query-result uniformity. For example, in [Smyth *et al.*, 2004] we have analysed different types of search behaviour by looking at specialised search tasks and groups of searchers to find a remarkable consistency in their individual search patterns. Thus in practice, we have found collaborative search to be capable of significantly improving the quality of search results.

The main focus of this paper and section is to consider how multiple communities of searchers might cooperate and collaborate during Web search. To this end we outline a straightforward adaptation of the collaborative Web search concept, which describes how groups of related communities can contribute potentially relevant results to the result-list of a host community. Thus, when a new target query, q_T, is submitted by a member of some host community, C_h, a result-list is generated with reference to C_h's local case-base in the usual way, but in addition extra novel results are also promoted by a set of related communities, C_1,...,C_k, in the hope that these extra results may also be relevant.

3.1 Collaborating Communities

So far we have focused our attention on a single hit-matrix, corresponding to a single community of searchers. But of course collaborative Web search contemplates the creation of multiple hit-matrices each tuned to the preferences of a different community. This is equivalent to the availability of a collection of distributed case-bases, each focusing on the needs of a different community of searchers. And recent work in the distributed CBR literature suggests that it is worthwhile considering the knowledge contained in other similar case-bases when solving some problem in a related task context [Leake and Sooriamurthi, 2001; McGinty and Smyth, 2001; Prasad and Plaza, 1996].

Thus, we can consider the search cases of a related community when recommending results to a searcher from a given host community. If there are strong similarities between two communities then the results from a related community may very well be relevant, or even provide an alternative perspective, for the searcher. For example, a searcher from a community focused on European skiing might benefit from the results of a community focused on skiing in North America. Both communities will have similar query distributions but different result selections, and our European skier might benefit from information about North American resorts when planning a winter break. Both sets of recommendations can be presented to the searcher. For example, the primary results from the host (European skiing) community can be presented as the main results and supplemented by an alternative set of recommendations from the North American community. The advantage of maintaining this community separation is that the searcher can better understand the origin of the different result sets, and the community titles will help them to understand the different focus of alternative results. This could be viewed as a novel alternative to the type of result clustering that search engines like Vivissimo (www.vivissimo.com) do except that our clusters correspond to real communities of interest as opposed to a mechanical separation of results into topical clusters that may or may not correspond to meaningful groupings in practice; see Section 5. Of course, both sets of recommendations could be integrated to form one unified list of recommendations; see Section 3.3.

3.2 Community Similarity

For the above approach to be useful we must look to the recommendations of communities that are demonstrably similar to the host community. But what do we mean by community similarity and how might it be measured? There are many ways to look at the concept of community similarity. For example we might start by supposing that if two communities have similar query term distributions then they might reflect the interests of two similar communities of users. However this is not necessarily the case, and not sufficient for our needs. For instance, a motoring community might share many queries with a community about wild cats (e.g., 'jaguar', 'puma', 'cougar' are all common car names) but very different result selections will have been recorded.

Thus, it makes more sense to look at the results that have been selected in response to searches as an estimate of community similarity. In this paper we start with a model of community similarity that is based on the proportion of overlapping results (see Equation 8 between a host community, C_h, and another community, C_r).

$$CommSim(C_h, C_r) = \frac{|Results(C_h) \cap Results(C_r)|}{|Results(C_h) \cup Results(C_r)|} \tag{8}$$

This similarity metric could be improved, it does not take into account any relationship between the query distributions of the communities. Ideally, it makes sense to consider community similarity as a function of both result overlap and query overlap. Moreover, the simple result overlap function above does not cater for the fact that two communities might share few result pages in common, but there may be strong

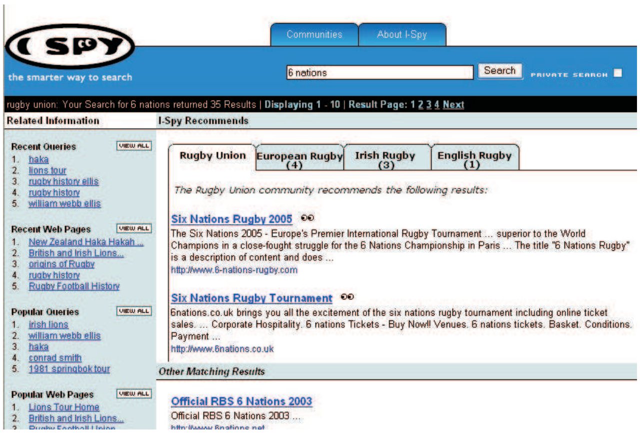

Figure 2: Rugby Union community recommended results.

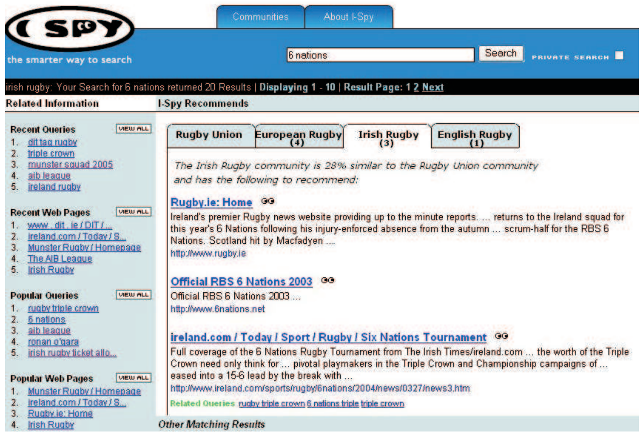

Figure 3: Irish Rugby community recommended results.

similarities between the contents of their pages. We will leave these enhancements as future work and proceed with the current model of community similarity for now.

3.3 Result Clustering & Ranking

Once a set of related communities has been identified (by their similarity to the host community) they can each be used to produce a set of results in response to the target query, q_T. In this case, for each related community we only seek to retrieve the set of result recommendations coming from their respective hit-matrices (search case-bases). Thus, each related community, C_i, produces a set of recommended results, R_i. These result-lists complement the result-list R_T that is produced for the host community; remember that R_T contains the promotions from the host community's hit-matrix *plus* those results that are returned from the underlying search engines in respond to q_T; see Section 2. As mentioned above there are two ways in which these different result-lists might be presented to the user, either as separate lists, with separate recommendations under their respective community headings, or as a combined list of recommendations. In this paper we focus on the benefits of the former, arguing that preserving the association between related communities and their recommended results provides the searcher with useful context information when it comes to understanding these recommendations. An example of this approach is presented in Figures 2 and 3 for a collection of search communities related to the sport of Rugby; these examples use the I-SPY system (ispy.ucd.ie) which is a robust, fully deployed version of collaborative Web search that currently uses Google and HotBot as its underlying search engines. The target query, '6 nations' (referring to the annual competition), is provided by a member of the host community, *Rugby Union*, and this community's results and recommendations are shown in Figure 2. Notice that in this screenshot there are two main sections of results. The top-half of the result list (labelled as *'I-Spy Recommends'* presents the recommendations from the host community; those results that have been selected and ranked from previous similar cases for this community. There are 2 such recommendations in this example. The second half of the result-list (labelled as *'Other Matching Results'* presents the remaining results that have come from the underlying search engines used by I-SPY, sorted by their meta-score; for reasons of space only one of these can be shown in Figure 2.

Notice, along the top of the recommended results, a set of tabs. Each containing the title of a related community followed by the number of recommended results found by this community. In this example, 4 related communities are shown, in order of their similarity to the host. Figure 3 presents the recommendations of the *Irish Rugby* community. Notice how these recommendations differ from those of the host community, although still clearly relevant to the target query. Also 2 of the 3 recommendations have an Irish angle; one is from an Irish rugby site (*Rugby.ie*) the other is an article from *Ireland.com* the Web site for an Irish newspaper. Before concluding it is worth commenting on the alternative presentation option whereby the results of these related communities are combined with those of the host to form a single recommendation list. To do this we need a way to score the recommendations from the related communities. Obviously they all have their own relevance scores, computed with respect to their own communities according to Equation 7. However these relevance scores can not be used in the context of the host community. Instead we propose that these relevance scores be modified to take account of the fact that they come from a related community. Specifically, each recommended result, p_j, from community, C_r, is discounted in proportion to the similarity between the host community (C_h) and the related community, as shown in Equation 9; note $c_1^r, ..., c_n^r$ is the set of search cases from C_r that are related to q_T. For reasons of space we will not pursue this approach any further here but will leave it as an option for future work.

$$CRel(C_h, C_r, p_j, q_T) = CommSim(C_h, C_r) \bullet WRel(p_j, q_T, c_1^r, ..., c_n^r) \quad (9)$$

4 Evaluation

In previous work we have demonstrated the benefits of collaborative Web search, through a range of live-user trials. For example, in [Smyth *et al.*, 2004; 2005] we present the results

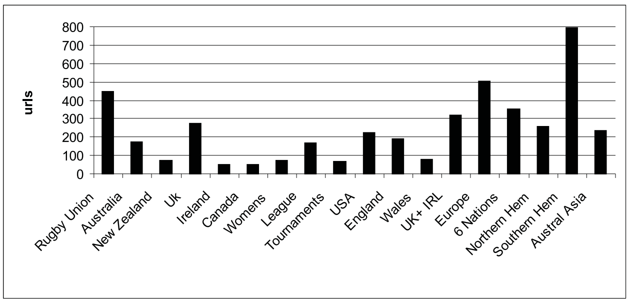

Figure 4: Number of URLs in each rugby community.

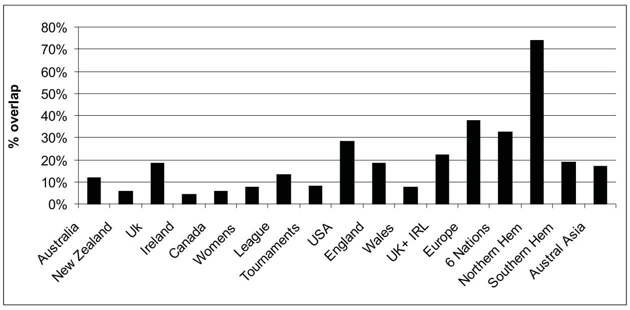

Figure 5: Similarity of the 16 communities compared to the Rugby Union community.

of two different user trials that show how collaborative Web search can significantly improve the precision and recall of an underlying search engine (Google in this case) with respect to the needs of a community of like-minded searchers.

In this paper we have speculated about the value of including recommendations from other search communities when responding to a query submitted in a specific host community. In particular, we have claimed that similar communities will recommend results that are related to the target query and the searcher's needs. Indeed we believe that, in general, communities that are more related to the host will be a more reliable source of relevant results. We also suspect that communities which are less closely aligned to the host may still have a role to play in suggesting results that are partially relevant and that might not otherwise be promoted by the host community. In this section we will describe the results of an experiment designed to test these hypotheses.

4.1 Methodology

This experiment was conducted using a mixture of artificial and real search data in order to simulate a society of related search communities. We constructed a range of topical search communities, simulated the activity of interested searchers within these communities, and then tested the quality of community recommendations using live-user search queries.

Community Creation

First we needed to generate a set of topical communities. To do this we used the subcategory listings from the Rugby domain within the Yahoo and ODP directories, generating communities that ranged from general communities such as *Rugby Union* to more specialised communities such as *Irish Rugby* and *6 Nations Championship*. The directory listings provided us with lists of Web pages to seed our communities. These lists were expanded using Google's 'similar page' facility; each seed was used to generate 10 similar pages which were validated and added to our communities as additional *relevant* URLs. In total 17 communities were created ranging in size from 45 URLs (for the Irish Rugby community) to 448 URLs (for the Rugby Union community); see Figure 4.

Populating Community Hit Matrices

Next we needed to generate each community search case-base by populating the hit-matrices from the behaviour of a set of simulated searchers. To do this we needed a set of plausible search queries and a way to simulate the selections of community members over the result-lists produced for these queries. For each community we generated a set of queries from the titles of its relevant pages; we constructed 10 queries per relevant page with an average of 2.39 terms per query. To simulate search, and train the hit-matrix, we submitted each query to HotBot and retrieved the top 500 results. We assumed that a community member would be interested in any HotBot result that was also one of the relevant pages for the community in question and the hit-matrix was updated as if the community member had selected this result.

Relevance Testing

Once the training phase was completed we had a set of community-based I-SPY search engines that we could test on a variety of novel test queries. Our test queries were real search queries extracted from the search logs of a rugby-specific search engine. This provided 393 test queries with an average of 2.47 terms per query. The queries included a range of general (e.g. rugby boots) and more specific (e.g. wanganui club rugby) queries, although a small number of non-rugby queries were also present.

During the testing phase we selected the *Rugby Union* community as the host community and measured its similarity to the 16 other communities; Figure 5 shows these similarity values. We replayed each query through I-SPY, just as if a searcher had entered the queries from the Rugby Union community. In parallel the queries were also replayed through all other communities. For each result-list we noted any result that was judged to be relevant to the host *Rugby Union* community. To be considered relevant to the *Rugby Union* community we availed, once again, of the Google 'similar pages' facility in a similar way to the manner in which we populated the communities with URLs. Essentially we extracted the top 50 similar pages for each of the URLs in the *Rugby Union* community, to produce a set of just over 1600 unique *Rugby Union* URLs. Any result that was a member of this expanded set was considered to be relevant to the *Rugby Union* community. Remember also that all of these communities had been trained on a different set of queries from the test queries.

4.2 Successful Searches

Perhaps the simplest measure of performance is to consider the percentage of queries for which the host and related com-

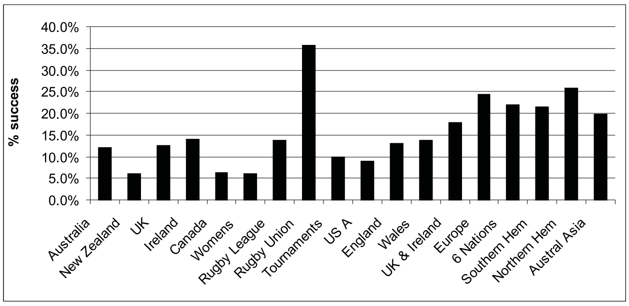

Figure 6: Success rate of each community

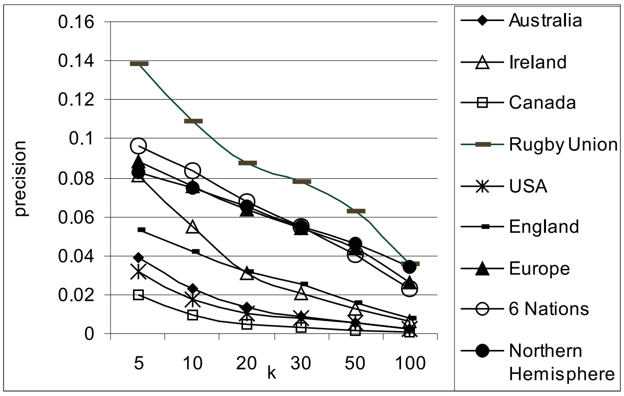

Figure 7: Precision of result lists

munities are able to retrieve at least one result that is considered relevant to the host; remember, we are concerned only with relevance to the host because our searcher is searching from this community. The results for each community are presented in Figure 6. They indicate that the *Rugby Union* community performs best overall (35% successful sessions), but a range of other communities also achieve comparable success scores. For example, communities such as *Europe* and *Northern Hemisphere* achieve success rates of about 25%. These communities are also the ones that have the greatest similarity with the host. Less similar communities like *Canada*, *Rugby League* and *Womens* perform less well, with success rates at about the 5% level.

It is worth noting that these results tell us that for 65% of the queries, the host community does not retrieve any relevant result. Keep in mind that our notion of relevance is an especially strong one in this experiment — there are only just over 1600 URLs that are considered to be *Rugby Union* — and that not all of the test queries are directly *Rugby Union* related. Results that are Rugby results in general but not Rugby Union in particular are not considered to be relevant. Nevertheless this means that there were 260 queries for which no core Rugby Union related results were retrieved by the host. However, when we look at the results from the other communities we find that 12.5% of these unsuccessful queries can be answered by the recommendations from at least one related community. In other words, often these related communities recommend results that are also returned by the host, but sometimes they recommend new results that are not returned by the host but that are relevant nonetheless; we will return to this issue at a finer level of detail in what follows.

4.3 Result Precision

The percentage of successful sessions is only a crude measure of search performance. In this section we consider the more traditional measure of precision – the percentage of retrieved results that are relevant – for different result-list sizes, $k = 5..100$. Figure 7 shows the precision scores for a selection of communities; presenting all 17 precision graphs is not feasible in the space available so we have selected a representative sample. As expected precision falls off as k increases; most relevant results appear high up in the result lists. Also, as expected, the precision for the host, *Rugby Union*, is significantly higher than all other related communities. For example, at $k = 5$ the average precision for *Rugby Union* is 0.14 whereas all other communities have a precision of less than 0.1; remember that precision for related communities is measured in terms of relevance to the host. This benefit is maintained as k increases. When analysing these scores we must remember that while all communities are rugby related each community represents a different aspect of the sport, many communities may not contain information relating to a query and so average precision scores drop.

We also see that there is a strong relationship between the precision characteristics of a related community and its similarity to the host. For example, the top 3 most similar communities (*N. Hemisphere, Europe, 6 Nations*) have the best precision characteristics (next to the host) and the least similar communities (e.g. *Canada*) have the poorest precision characteristics. In fact the correlation coefficient between the average precision of each community (across all values of k) and the similarity of these communities to the host is 0.72, indicating a strong dependency between community similarity and the relevance of its recommendations to the host.

4.4 Result Uniqueness

The results in Figure 7 tell us that similar communities to the host can make relevant recommendations, a feature that is certainly important in situations where the host cannot make strong recommendations of its own. This can occur, for example, when the host community is immature and has not had an opportunity to adapt to its members. However, even when the host community can make strong recommendations of its own it may still make sense to look to the recommendations of related communities, especially if they are capable of making new, distinctive recommendations. The results presented in Figure 8 show the percentage of relevant recommendations made by a related community that are not made by the host, for different values of k. For example, we see that 70-80% of the relevant recommendations made by the *USA* community are unique with respect to the host community's recommendations. These recommendations offer our *Rugby Union* searcher distinctive results that are relevant but that are not found in the host result-list and that, therefore, would have been missed under normal circumstances. Indeed, in gen-

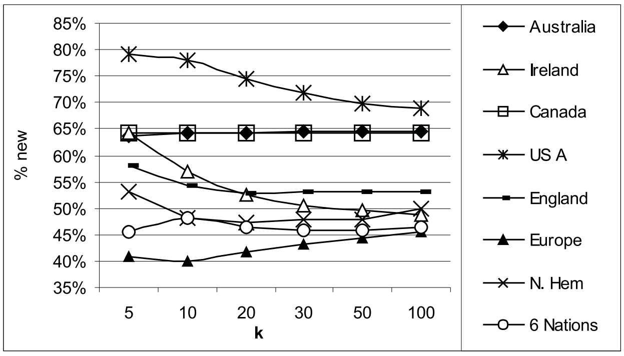

Figure 8: Percentage of new relevant results that were not in the Rugby Union result lists

eral, we see that those communities that are less similar to the host are the ones that are making more of these relevant new recommendations, indicating an important trade-off between community similarity, on the one hand, and recommendation precision and distinctiveness on the other.

4.5 Summary

This evaluation was undertaken to determine whether search communities other than the host can serve as a viable sources of alternative recommendations. The results indicate that they can. They show that similar communities are capable of making recommendations that are potentially very relevant to the host. And they also show that even moderately similar communities have an important role to play when it comes to recommending distinctive results that may be missed by the host.

Of course this evaluation is not perfect. It relies, in part at least on artificial search data; however it also uses live-user search data. Ultimately, the value of the results will depend critically on how we have used the artificial user data. In this case, we have used artificial data to build our search communities, by seeding the communities with pages that are known to be relevant to a topic, and then simulating the search behaviour of users of these communities by using an artificial set of queries drawn from the terms in the community pages. How confident can we be that this will produce a reasonable set of realistic communities and hit-matrices? Of course this needs to be validated with live user trials to fully answer this question. However, we would point out that we have used similar methodologies in the past [Freyne *et al.*, 2004], during earlier evaluations of collaborative Web search, and that these earlier results have proved to be consistent with our recent live-user trials. Hence, we argue that this use of artificial user data is likely to be reasonable. We can be similarly confident about appropriateness of our testing methodology, where genuine user queries, drawn from real search logs, are used.

In summary then, our evaluation results highlight the *potential* benefits of cooperation among communities in collaborative Web search. Obviously, we will move towards larger-scale live user trials in the future, but the fact that our current results point to a clear benefit for collaborative Web search bodes well for these future evaluations.

5 Conclusions and Related Work

The work presented touches on a number of areas of related research by combining ideas from Web information retrieval and case-based reasoning. Of particular importance is the idea that Web search experience can be usefully captured as a case-base of reusable cases and that this experience can be distributed across multiple case-bases which correspond to the different needs of different communities of searchers.

Of course there is a long history of the use of case-based methods in a variety of information retrieval tasks. For example, the work of Rissland [Rissland and Daniels, 1995] looks at the application of CBR to legal information retrieval (see also [Ashley, 1990]), and Burke et al. [Burke *et al.*, 1997] describe a case-based approach to question-answering tasks. However these approaches have all tended to focus on particular application domains rather than the broader area of Web search; see also [Lenz and Ashley, 1999]. That said there is some CBR work in the broader context of Web search. For example, the *Broadway* recommender system [Kanawati *et al.*, 1999] is notable for its use of case-based techniques to recommend search query refinements, based on refinements that have worked well in the past. Perhaps more related to the core work in this paper is the *PersonalSearcher* [Godoy and Amandi, 2000] which combines user profiling and textual case-based reasoning to dynamically filter Web documents according to a user's learned preferences.

The idea that experience (in this case, search experience) can be distributed across multiple case-bases is not new, and in recent years many researchers have considered the use of multiple case-bases during problem solving; see [Leake and Sooriamurthi, 2001; McGinty and Smyth, 2001; Prasad and Plaza, 1996]. For example, Leake et al. [Leake and Sooriamurthi, 2001] consider the benefits and challenges when reusing the experience of multiple case-bases that reflect different tasks and environments. They look at prediction problems and consider how a local case-base can usefully determine when to look to external case-bases as a source of knowledge, and how external cases might be adapted in line with the local task and environment; see also the work of [McGinty and Smyth, 2001] for ideas in the planning domain.

Finally it is worth commenting on recent result clustering work in Web search that is also related to the work of this paper. The ranked-list presentation format that has been almost universally adopted by most Web search engines makes it quite inefficient for users to quickly assess the relevance of retrieved results. One solution that has attracted some considerable attention in recent years is results clustering. The basic idea is to organise result-lists into clusters of semantically related results [Zamir and Etzioni, 1999] and a range of efficient clustering algorithms have been developed for the rapid clustering of search results [Dell Zhang, 2004; Hamilton, 2003]. The traditional approach to result clustering involves an analysis of the contents of result pages, or their associated snippet texts. This is computationally expensive and accurate clustering is often compromised for reasons of response-time. Our extension to collaborative Web search is a form of result clustering, in the sense that each community contributes as a cluster of results, but without the need for any result-based analysis. Moreover, the clusters relate to genuine groups of interest, which may improve their quality and interpretability, although this has not been tested as yet.

In summary, we have shown how Web search experience can be captured at community level as a case-base of search cases, and how this experience can be distributed across multiple case-bases. In addition, the novel aspect of this work has been to consider how these distributed case-bases might be reused by different host communities. To this end we have shown that the experience of communities of searchers can be usefully leveraged to help searchers from other communities. These related communities can serve as a source of recommendations that are both relevant and distinctive.

References

[Aamodt and Plaza, 1994] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1):39 – 59, 1994.

[Ashley, 1990] K.D. Ashley. Modeling legal argument: Reasoning with cases and hypotheticals. Technical 88–01, Department of Computer and Information Science, University of Massachusetts., Amherst, MA, 1990.

[Balfe and Smyth, 2005] E. Balfe and B. Smyth. A Comparitive Analysis of Query Similarity Metrics for Community Based Web Search. In *Proceedings of the 6th International Conference on Case-Based Reasoning*, 2005.

[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.

[Burke *et al.*, 1997] Robin Burke, Kristian Hammond, Vladimir Kulyukin, Steven Lytinen, Noriko Tomuro, and Scott Schoenberg. Question answering from frequently asked question files: Experiences with the faq finder system. *AI Magazine*, 18(2):57–66, 1997.

[Dell Zhang, 2004] Yisheng Dong Dell Zhang. Semantic, Hierarchical, Online Clustering of Web Search Results. In *In Proceedings of the 6th Asia Pacific Web Conference (APWEB).*, Hangzhou, China, 2004.

[Freyne *et al.*, 2004] J. Freyne, B. Smyth, M. Coyle, E. Balfe, and P. Briggs. Further Experiments on Collaborative Ranking in Community-Based Web Search. *Artificial Intelligence Review: An International Science and Engineering Journal*, 21 (3-4):229–252, June 2004.

[Furnas *et al.*, 1987] George W. Furnas, Thomas K. Landauer, Louis M. Gomez, and Susan T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):964–971, 1987.

[Godoy and Amandi, 2000] Daniela Godoy and Anala Amandi. PersonalSearcher: An Intelligent Agent for Searching Web Pages. In *IBERAMIA-SBIA*, volume 1952, pages 62–72. Springer, 2000.

[Hamilton, 2003] N. Hamilton. The mechanics of a deep net metasearch engine. In *Proceedings of the 12th International World Wide Web Conference*, Budapest, Hungary, 2003.

[Kanawati *et al.*, 1999] R. Kanawati, M. Jaczynski, B. Trousse, and J-M Andreoli. Applying the broadway recommendation computation approach for implementing a query refinement service in the cbkb meta-search engine. In *Coinference Fraicaise sur le Raisonnement a Partir de Cas (RaPC'99)*, 1999.

[Kleinberg, 1999] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[Lawrence and Giles, 1998] S. Lawrence and C. Lee Giles. Context and Page Analysis for Improved Web Search. *IEEE Internet Computing*, July-August:38–46, 1998.

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

[Leake and Sooriamurthi, 2001] David B. Leake and Raja Sooriamurthi. When two case bases are better than one: Exploiting multiple case-bases. In *Proceedings of the 4th International Conference on Case-Based Reasoning.* Springer-Verlag, 2001.

[Lenz and Ashley, 1999] M. Lenz and K. Ashley. Aaai workshop on textual case-based reasoning. Technical Report ws-98-12, AAAI, 1999.

[McGinty and Smyth, 2001] L. McGinty and B. Smyth. Collaborative case-based reasoning: Applications in personalized route planning. In *Proceedings of the 4th International Conference on Case-Based Reasoning.* Springer-Verlag, 2001.

[Prasad and Plaza, 1996] M.V. Nagendra Prasad and E. Plaza. Corporate memories as distributed case libraries. In *Proceedings of the Corporate Memory and Enterprise Modeling Track in the 10th Knowledge Acquisition Workshop*, 1996.

[Rissland and Daniels, 1995] Edwina L. Rissland and Jody J. Daniels. A hybrid cbr-ir approach to legal information retrieval. In *ICAIL '95: Proceedings of the fifth international conference on Artificial intelligence and law*, pages 52–61. ACM Press, 1995.

[Scott, 1991] J. Scott. *Social network analysis: A handbook.* Sage Publications, CA, 1991.

[Smyth *et al.*, 2004] B. Smyth, E. Balfe, J. Freyne, P. Briggs, M. Coyle, and O. Boydell. Exploiting query repetition and regularity in an adaptive community-based web search engine. *User Modeling and User-Adaptated Interaction*, 14(5):383–423, 2004.

[Smyth *et al.*, 2005] B. Smyth, M. Coyle, O. Boydell, P. Briggs, E. Balfe, J. Freyne, and K. Bradley. A Live-User Evaluation of Collaborative Web Search. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI-05*, 2005. Edinburgh, Scotland.

[Zamir and Etzioni, 1999] Oren Zamir and Oren Etzioni. Grouper: Dynamic clustering interface to web search results. *Computer Networks*, 31(11-16):1361–1374, 1999.