
Volume 46, issue 6, November 2010
ISSN 0306-4573
ELSEVIER
INFORMATION PROCESSING & MANAGEMENT
An International Journal
Email: ipm@unisi.ch
Web site: http://www.elsevier.com/locate/infoproman
Editor-in-Chief: Fabio Crestani, University of Lugano
Associate Editors: Ricardo Baeza-Yates, Yahoo! Research Barcelona
Kalervo Järvelin, University of Tampere
Noriko Kando, National Institute of Informatics
Efthimis N. Efthimiadis, University of Washington
Chengxiang Zhai, University of Illinois at Urbana-Champaign
SPECIAL ISSUE: Collaborative Information Seeking
Guest Editors: Gene Golovchinsky, Meredith Ringel Morris and Jeremy Pickens

Information Processing and Management 46 (2010) 773–781

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Role-based results redistribution for collaborative information retrieval

Chirag Shah [a,*], Jeremy Pickens [b], Gene Golovchinsky [b]

[a] School of Information & Library Science (SILS), University of North Carolina at Chapel Hill NC 27599, United States
[b] FX Palo Alto Laboratory (FXPAL), Calironia, United States

ARTICLE INFO

Article history:
Received 16 May 2009
Received in revised form 22 September 2009
Accepted 2 October 2009
Available online 4 November 2009

Keywords:
Collaborative search
Algorithmic mediation
Evaluation

ABSTRACT

We describe a new approach for algorithmic mediation of a collaborative search process. Unlike most approaches to collaborative IR, we are designing systems that mediate explicitly-defined synchronous collaboration among small groups of searchers with a shared information need. Such functionality is provided by first obtaining different rank-lists based on searchers' queries, fusing these rank-lists, and then splitting the combined list to distribute documents among collaborators according to their roles. For the work reported here, we consider the case of two people collaborating on a search. We assign them roles of Gatherer and Surveyor: the Gatherer is tasked with exploring highly promising information on a given topic, and the Surveyor is tasked with digging further to explore more diverse information. We demonstrate how our technique provides the Gatherer with high-precision results, and the Surveyor with information that is high in entropy.

1. Introduction

It is natural for humans to collaborate on difficult tasks (Denning & Yaholkovsky, 2008), including exploring and retrieving information (Morris, 2008). Twidale, Nichols, and Paice (1997) argued that introducing support for collaboration into information retrieval systems would help searchers to learn and use the systems more effectively. They further claimed that a truly user-centered system must acknowledge and support collaborative interactions between people, and showed that searchers often desire to collaborate on search tasks. Based on their extensive study of patent office workers, Hansen and Jarvelin (2005) also concluded that the assumption that information retrieval performance is purely individual needs to be reconsidered. While this issue of collaboration has attained considerable attention lately, the mechanics of how to mediate such collaboration remain largely unexplored. Joho, Hannah, and Jose (2009) identified three conceptual approaches to facilitate such collaboration in information searching tasks: models, techniques, and interfaces. In this paper, we propose a new model for addressing role-based IR in collaboration. The model uses existing techniques for merging and splitting results in a unique way.

We propose to address a core issue of facilitating collaboration between a pair of searchers via algorithmic mediation. More specifically, we describe the design of a system to mediate explicit, synchronous collaboration (Golovchinsky et al., 2009) between multiple searchers. Mediated collaboration is typically not necessary for simple precision-oriented searches such as fact-finding or for known-item search. However, for tasks that are more recall-oriented, more complex and exploratory in nature, and for situations in which the information need is not satisfied by a single document, there is value in having a system that explicitly supports multiple, task-aligned searchers. In addition to this, we wish to incorporate two major elements of algorithmically mediated collaboration as expressed by Foley (2008): sharing of knowledge, and division of labor. The former is achieved by supporting two different roles of the participants, and the latter is done by splitting the returned results of a search.

* Corresponding author.
E-mail addresses: chirag@unc.edu (C. Shah), jeremy@fxpal.com (J. Pickens), gene@fxpal.com (G. Golovchinsky).

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2009.10.002

While we recognize the importance of interactivity within collaborative systems, in this paper we focus on one round of a search process. Studying the effects of single-round interaction should aid future interactive system design. Focusing on a single round also allows us to specifically study the effects of our algorithmic decisions; future work will address how to incorporate these techniques into an iterative search process.

One way to structure a collaborative retrieval system is to identify roles of the collaborators, and to describe the interfaces in terms of supporting these roles (Pickens, Golovchinsky, Shah, Qvarfordt, & Back, 2008). In this paper, we define two complementary roles – Gatherer and Surveyor – for our team. The goal of the Gatherer is to scan results of the joint search activity of team members to discover the most immediately relevant information. The goal of the Surveyor is to browse a wider diversity of information to get a better understanding of the nature of the collection being searched, to understand where the current queries might be failing, and to identify potential avenues of exploration. While each person contributes to the querying process by expressing their individual understanding of the information need through their queries, the results they see are affected by their roles. The "black box" facilitates this process by combining the ranked lists from the users' query formulations, and then by splitting the combined list based on role criteria. The goal of results combination is to improve overall precision by aligning or combining multiple users' query formulations. The objective of splitting, on the other hand, is to preserve that improvement while simultaneously encouraging exploration and information diversity to avoid getting trapped in a rut.

It is important to note that these components are not intended to improve the performance of the basic retrieval algorithm. They are merely acting as each other's complements to facilitate algorithmic mediation for role-based collaboration. Our primary evaluation, thus, will be based on comparing the inputs and outputs of this "black box." In addition, we will also analyze merging and splitting phases, testing various retrieval parameters with a set of methods. The main contribution of our work is the demonstration of an effective way of using these two complementary processes to create a system that distributes search results to team members in a role-based collaborative IR session.

2. Background

Terms such as *collaborative IR, co-browsing,* and *social search* have been used in the literature to label similar approaches to information retrieval. We will first review some of these approaches to put our work in the context. In an IR system, there are various objects and processes that could be combined in "collaborative" fashion. For instance, Fu, Kelly, and Shah (2007) showed how different queries from a set of users for the same information goal can be combined, calling it "collaborative queries", for better retrieval performance. Rather than simply combining queries, we first combine and then distribute the rank-lists produced by executing the individual queries.

The term "collaboration" is used frequently to describe recommender systems, which can be classified into three main categories: content-based, collaborative, and hybrid recommendation approaches (combination of the above two) (Adomavicius & Tuzhilin, 2005). Content-based recommender systems rely on similarity of content selected by earlier users to make new recommendations. Collaborative recommender systems user other people's opinions or behaviors to inform a new person's search process. Such systems are also called collaborative filtering systems. The majority of collaborative systems are driven by the motivation that prior people's behavior can inform a new user (Linden, Smith, & York, 2003), and that a group of users can tackle a complex problem better than any one of them individually (Denning, 2007). Smyth, Balfe, Briggs, Coyle, and Freyne (2003) argued that one way of making it possible to connect people to information that is difficult to find is to incorporate collaboration in the search phase of an information seeking process. They showed how collaborative search can act as a front-end for existing search engines and re-rank results based on the learned preferences of a community of users. They demonstrated this concept by implementing I-Spy system (Freyne, Smyth, Coyle, Balfe, & Briggs, 2004). I-Spy, however, acts more like collaborative filtering than like synchronous collaborative searching, which is the focus of our work. HeyStaks[1] supports multiple kinds of interaction, along the lines of SearchTogether (Morris & Horvitz, 2007). Both systems allow users to contribute documents independently, but the system synchronizes the data so that the latest documents are available to all people using a given Stak. While HeyStak may re-order search results by promoting previously-selected documents, it does not distinguish between different roles that searchers may have. It may be thought of as a light-weight recommendation overlay on top of the regular search engine ranking algorithm.

Another example is *AntWorld* (Menkov, Neu, & Shi, 2000), a tool developed to make it easier for the members of a common interest user group to collaborate in searching the Web. *AntWorld* harnesses the expertise of members of a common interest group through their evaluation of documents encountered while searching. It uses judgments about documents made by some people to guide others to pages they may find useful. Once again, this serves more as a collaborative filtering system than a synchronous collaborative search system. While our work incorporates multiple users sharing information and affecting each other by their actions, such sharing and influence are mediated by the system in real time, so that all team members benefit from each other's current (and past) activities. Also, in contrast to AntWorld, the information seen by different users of the system is intentionally kept different, allowing the collaborators work on different aspects of the same task.

[1] http://www.heystaks.com

It is important to note here that in many of these recommmendation-driven applications, a person receiving the recommendations may not know the other users in the network personally, and may not share the same information need. Thus, a user is not intentionally and interactively engaged in a true collaboration with other users; he is merely getting filtered content based on other users' actions on similar information. There have been some applications that exploit more tightly connected social networks instead of the entire network to filter and recommend information. For instance, Kautz, Selman, and Shah (1997) described *ReferralWeb*, which was based on providing recommendations via chains of named entities instead of anonymous users in the network. Even in this case, however, people searching for information could not be certain that others used the system in sufficiently similar ways for their prior experiences and actions to be useful.

A common way of providing a collaborative IR solution is to extend a traditional IR model to incorporate multiple people. For instance, Kuhlthau (1991) described the Information Search Process (ISP) model to explain information seeking from a searcher's perspective. She later extended this model to bridge the gap between information seeking and retrieval (Kuhlthau, 2005). This model inspired Hyldegárd (2006) to study information seeking and retrieval process in a group-based educational setting. Our work differs in that we are not simply extending a typical IR model to include multiple searchers, but instead, we propose a new technique that can leverage the diversity of people's skills and experience when working toward a shared information need.

A typical use of a collaborative searching or browsing facility is to let a group of people share information. For instance, Root (1988) introduced the idea of social browsing to support distributed cooperative work with unplanned and informal social interaction. This idea was carried over by Donath and Robertson (1994) several years later as *The Social Web* that allows a person to know that others were currently viewing the same web page and communicate with those people. Wittenburg, Das, Hill, and Stead (1995) came up with the notion of Group Asynchronous Browsing (GAB) to provide tools for people to leverage the information hunting and gathering activities of other people or group of people on the Web. Cabri, Leonardi, and Zambonelli (1999) showed how cooperative Web browsing can be supported using a proxy, and Gerosa, Giordani, Ronchetti, Soller, and Stevens (2004) presented a similar idea with the application of e-learning. Systems such as *GroupWeb* (Greenberg & Roseman, 1996), *GroupScape* (Graham, 1997), and *Co-Vitesse* (Laurillau & Nigay, 2002) are meant to provide a co-browsing interface to allow users in a collaborative environment to explore the Web together. Keller, Wolfe, Chen, Rabinowitz, and Mathe (1997) described *WebTagger*, a social bookmarking service similar to del.icio.us.[2]

The work presented in this paper differs from other systems listed above in that we provide algorithmic mediation to an explicitly established collaboration among users, rather than only shared awareness. In the literature we find several approaches that entertain such explicit collaboration. For instance, Twidale and Nichols (1996) described the *Ariadne* system that allowed a user (patron) to collaborate with an information expert (reference librarian) remotely and synchronously. However, the distribution of the information was handled manually, by the users, without much help from the system.

SearchTogether (Morris & Horvitz, 2007) and Shah, Marchionini, and Kelly (2009) allow people to contribute queries and share search results. They are good examples of UI-level mediation for explicit collaborative search (Golovchinsky et al., xxxx). Such systems have been used as platforms to investigate various aspects of UI-level mediation, including awareness and sense-making (Paul & Morris, 2009). SearchTogether has also recently introduced algorithmic multi-user results allocation ("smart splitting") (Morris, Teevan, & Bush, 2008). That work uses personal information such as documents and browsing history to cluster search results from single queries for splitting among users; personalization meets division of labor. Searchers get to evaluate the documents they personally are most likely to recognize as useful, as extracted from a single expression of an information need. Our approach differs from this in several ways. We begin by pooling independently-entered queries to create a richer set of results. A more important difference is at the second, splitting stage. Rather than splitting of results based on personal traits of the searchers, we split results based on role-oriented sub-tasks or goals. The focus on (and optimization for) the role, rather than for the person, is an important distinction for future system and algorithm design. Other work on explicit collaborative search includes desktop systems such as Físchlár-DiamondTouch (Smeaton, Lee, Foley, McGivney, & Gurrin, 2006) which implement symmetric roles: users contribute search terms to a single query, every user sees the same result set at the same time, and each can grab results from that query for analysis.

We are interested in exploring mechanisms for distributing search results to people based on their roles. We are aware of very few systems that experiment with asymmetric roles in information seeking. The work described here builds on our earlier work (Pickens et al., 2008) that showed how one user in a collaboration could specialize in issuing queries and the other could specialize in evaluating the results. The work reported here is different in the way we define the roles, and in how we assign documents to each role. Instead of task-based roles, here we have information content-based roles. In other words, instead of letting each person be responsible for a certain kind of task (issuing query vs. looking at the results), we have defined the roles such that both people do the same tasks, but are shown information that is based on their roles (Gatherer or Surveyor). This instance of division of labor (Foley, 2008) represents situations in which people have fundamentally similar skill sets, but may benefit from adopting different approaches to exploring the data. For instance, a one searcher (in the Surveyor role) may explore the information landscape broadly, looking for serendipitous discovery, while another (in the Gatherer role) pursues specific highly relevant information from a particular aspect. As the session progresses, they can trade off roles. In this manner, roles offer collaborators a method for solving a potentially tedious task that otherwise would not be possible.

[2] http://del.icio.us

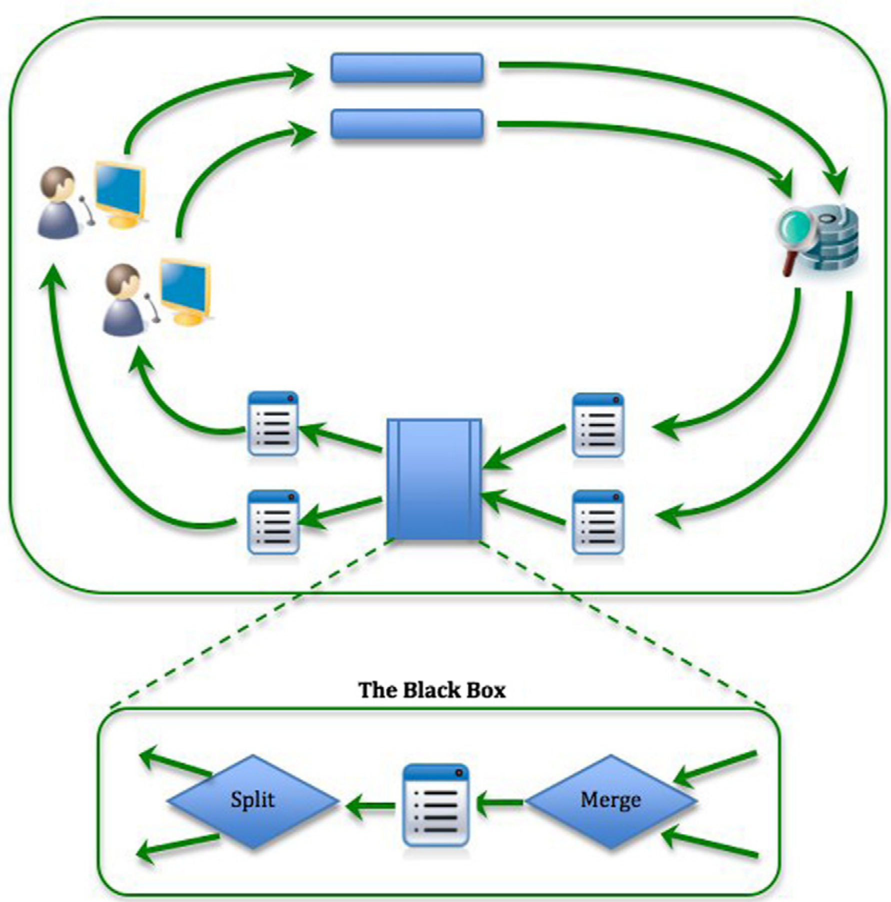

Fig. 1. Our proposal for merging and splitting rank-lists based on user roles.

Our proposal for a system that facilitates explicitly-defined synchronous collaboration between two collaborators with different roles is shown in Fig. 1. The system combines results of queries from two searchers and then redistributes them according to their roles. While there are several ways of defining these roles, as well as many possible algorithms for implementing them, in this work we examine the *Gatherer*, who focuses on quickly finding as much relevant information as possible, and the *Surveyor*, who will delve into a more diverse set of documents.

3. Role-based algorithmic collaboration mediation

In this section we describe our collaboration mediation algorithms. The collaborative scenario that we envision consists of three fundamental steps.

1. Two searchers are working together on an explicitly-shared information need. One searcher enters a query, and begins to review the resulting document set. Some time later, the second searcher enters a query.
2. The system evaluates the second searcher's query and displays it. The system also fuses the results of the two queries to produce a ranked list of higher precision than either query, individually.
3. The system then splits the fused results into two sublists, one for each person. These partitioned results are created based on the following role-based criteria:
   (a) One list is optimized for effectiveness (high precision), and presented to the first searcher – the Gatherer.
   (b) One list is optimized for exploration (high diversity), and presented to the second searcher – the Surveyor.

In essence, the first searcher gets the list that is most similar to the part of the list that he had already seen before the second query was available for fusion; the second searcher gets the other list. This assignment strategy should make recommendations in line with what has already been seen by the first searcher, making it easier to make sense of the additional suggestions, particularly if there is little overlap among query results. The suggestion process can be iterated as the system receives additional queries.

In the remainder of this section we will explain how the merging and splitting is done, and in the following section we will evaluate various aspects of the approach. We note that the main contribution of this paper is not in any particular merging or splitting algorithm, but in the overall conceptual scenario in which these algorithms are being used, i.e., to support intentional, synchronous collaboration of multiple users on a shared information need.

3.1. Merging

We use CombSUM fusion (Aslam & Montague, 2001) to merge the two users' queries. CombSUM first normalizes the raw scores for each document returned by a query, sums these normalized scores, and then re-ranks documents by the total sum. Documents originally ranked highly in multiple lists will receive a higher score than documents ranked highly in only a single list, while documents ranked lower, but found in multiple lists can end up ranked ahead of documents that were originally ranked higher, but only in a single list.

For our baseline comparisons we implemented a simple round robin merging approach. Round robin successively adds to the merged list the first two documents from each individual list, then the second two, and so on. If a particular document has already been added to the final merged list (e.g., because it was ranked higher in the other user's list), it is skipped and the next document is added instead.

3.2. Splitting

We split the combined list of documents using k-means clustering based on document content with k equal to the number of collaborators, which, in this case, is two. The idea is that searchers can cover different aspects of the search results, as appropriate to their roles. The Gatherer ensures that the most relevant documents are discovered as quickly as possible. The Surveyor scans other, more diverse documents, trying to understand the collection and to make serendipitous discoveries. In some sense this approach is similar to the scatter phase of Scatter–Gather, in which one cluster has typically been shown to have good recall and precision (Hearst & Pedersen, 1996).

Starting with the merged ranked list produced by merging query results from the two collaborating searchers, the top 2000 documents are randomly (but with equal probability) assigned to one of two clusters. The Vector Space centroid of each cluster is computed, and then documents are reassigned to each cluster based on nearest proximity to a centroid. Centroid calculation and reassignment is repeated for five iterations. Documents in each cluster are then presented to the two searchers in the same order they appear in the original merged list. While other clustering algorithms are possible, we believe this approach is representative.

4. Experiments

To create our test collection, we extracted terms from the description field of TREC topics 301–450, and ran these terms as queries to identify potential topics. We then selected all topics of moderate difficulty (precision@10 values between 0.1 and 0.5). This yielded 53 topics, from which we randomly removed three to produce five groups of 10 topics. For each group, we generated a paper form with topic descriptions and instructions asking people to write the query they would issue to a typical search engine for the given topic. Fifteen participants were asked to fill forms at their leisure, providing us one query for each of the 10 topics assigned to them, resulting in 150 queries for the 50 topics, or three queries per topic.

4.1. Merging

We used TREC relevance judgments for the selected topics the define the gold standard against which we compared query performance. To simulate single-iteration collaboration in group of searchers, we first ran the three queries (A, B, and C) for each topic individually. Next, we merged the ranked lists two ways (AB, AC, BC) using both round robin and CombSUM fusion. We then compared each of these merged lists against the individual query baseline, i.e., AB vs. A, AB vs. B, AC vs. A, AC vs. C, etc. This yielded a total of 300 comparisons from our initial 50 queries. The results are shown in Table 1. Overall, the round robin approach worked only slightly better than no merging, but the difference is not statistically significant. CombSUM, on the other hand, well out-performs the individual queries, and the results are significant.

Next, we compared CombSUM against round robin merging, directly. In this case there are only 150 comparisons to make – AB, AC and BC for 50 topics. In line with Shaw and Fox (1994), these results (Table 2) show that CombSUM fusion far outperforms round robin merging. This suggests that although ranked list fusion is not the only possible foundation for collaborative algorithms, it is a reasonable starting point.

4.2. Splitting

Now that we have combined the individual queries from multiple users, the next step in a single round of algorithmically mediated collaboration is to redistribute those results back out to the users. The most straightforward method for results redistribution is the "split", or de-round robin (de-RR) function, a dealing out of the results back to the users (Morris & Horvitz, 2007): One person would get the documents ranked 1, 3, 5, 7, 9, etc. from the combined list, and the other person would get results 2, 4, 6, 8, 10, etc. A stronger alternative is to use a clustering algorithm described in Section 3.2. By examining the

Table 1
Round robin merging and CombSUM fusion compared against an individual (non-merged) query baseline.

	Individual	Round robin	Chg (%)	*t*-Test	CombSUM	Chg (%)	*t*-Test
Total number of documents over all queries							
Retrieved	300,000	300,000			300,000		
Relevant	27,378	27,378			27,378		
Rel.|ret	7948	8646	+8.78	0.3085	9148	+15.10	0.0872
Precision							
At 5 docs	0.3133	0.3147	+0.4	0.9549	0.3693	+17.9	0.0243*
At 10 docs	0.2707	0.2800	+3.4	0.6368	0.3100	+14.5	0.0523
At 15 docs	0.2391	0.2418	+1.1	0.8759	0.2729	+14.1	0.0552
At 20 docs	0.2157	0.2240	+3.9	0.5770	0.2457	+13.9	0.0523
At 30 docs	0.1856	0.1922	+3.6	0.6062	0.2127	+14.6	0.0459*
At 100 docs	0.1078	0.1113	+3.3	0.6521	0.1208	+12.1	0.1005
At 200 docs	0.0734	0.0759	+3.4	0.6584	0.0831	+13.1	0.0920
At 500 docs	0.0427	0.0452	+6.0	0.4693	0.0483	+13.2	0.1165
At 1000 docs	0.0265	0.0288	+8.8	0.3085	0.0305	+15.1	0.0872
Average precision (non-interpolated)							
	0.0931	0.0931	–0.09	0.9918	0.1086	+16.63	0.0922

* Indicates *t*-test significance at $p < 0.05$.

Table 2
CombSUM fusion compared directly against Round robin merging.

	Round robin	CombSUM	Chg (%)	*t*-Test
Total number of documents over all queries				
Retrieved	150,000	150,000		
Relevant	13,689	13,689		
Rel.|ret	4323	4574	+5.81	0.0009*
Precision				
At 5 docs	0.3147	0.3693	+17.4	0.0002*
At 10 docs	0.2800	0.3100	+10.7	0.0115*
At 15 docs	0.2418	0.2729	+12.9	0.0005*
At 20 docs	0.2240	0.2457	+9.7	0.0015*
At 30 docs	0.1922	0.2127	+10.6	0.0000*
At 100 docs	0.1113	0.1208	+8.5	0.0003*
At 200 docs	0.0759	0.0831	+9.4	0.0001*
At 500 docs	0.0452	0.0483	+6.8	0.0004*
At 1000 docs	0.0288	0.0305	+5.8	0.0009*
Average precision (non-interpolated)				
	0.0931	0.1086	+16.74	0.0000*

* Indicates *t*-test significance at $p < 0.05$.

content of the retrieved results and dividing that content into two clusters, we can present each searcher with a separate cluster, one that is optimized for precision and the other optimized for diversity.

Table 3 reports the results of both de-RR and our proposed clustering approach, comparing with the original fused list from which each approach was constructed. For averaging purposes, we group each subset into the "best" and the "worst" subset, as measured by average precision.

As we can see, even though there is a slight improvement on precision at five documents, the best k-means cluster is essentially equivalent to the original fused list. On the other hand, the best de-RR cluster is significantly worse. This is to be expected, as there are a lot of relevant documents that will be split evenly between the two collaborators, thus lowering the precision of both. Indeed, when we look at the "worst" retrieval subsets, we can see that the k-means cluster is much worse than the de-RR cluster. This demonstrates that it is possible to re-distribute results in such a manner so as to preserve high precision for one of the users. The role of Gatherer becomes possible.

But what about diversity, or support for the Surveyor role? Here the goal is to satisfy the team's need to obtain a well-rounded view of the entire result set. When we re-examined the two clusters produced by the k-means step, we found that 112 times out of 150 (74%), the cluster with the lowest average precision was also the cluster with the highest average distance (or spread) from the cluster centroid (sign test $p \ll 0.01$). Although not a perfect measure of diversity, this indicates that while the Gatherer retains high average precision (AP), the Surveyor can simultaneously more quickly get a sense of the diversity of the documents in the results set. By way of comparison, the de-RR splitting of the same initial fused list does not exhibit these same role-based properties. With de-RR, we found that only 87 times out of 150 (57%, sign test $p < 0.10$) was

Table 3
Splitting of the CombSUM fusion list. Splits of this list are made via de-RR and via k-means clustering. The highest precision (best) clusters for each query are averaged together, and the lowest precision (worst) clusters for each query are also averaged together.

	CombSUM	Better average precision clusters				Lower average precision clusters			
		de-RR	Chg (%)	k-Means	Chg (%)	de-RR	Chg (%)	k-Means	Chg (%)
Total number of documents over all queries									
Retrieved	150,000	150,000		150,000		150,000		150,000	
Relevant	13,689	13,689		13,689		13,689		13,689	
Rel.|ret	4574	2952	−35.46*	4140	−9.49*	2509	−45.15*	1062	−76.78*
Precision									
At 5	0.3693	0.3867	+4.7	0.3907	+5.8	0.2333	−36.8*	0.1240	−66.4*
At 10	0.3100	0.3000	−3.2	0.3140	+1.3	0.1913	−38.3*	0.0987	−68.2*
At 15	0.2729	0.2564	−6.0	0.2711	−0.7	0.1689	−38.1*	0.0831	−69.5*
At 20	0.2457	0.2277	−7.3*	0.2523	+2.7	0.1523	−38.0*	0.0737	−70.0*
At 30	0.2127	0.1853	−12.9*	0.2151	+1.1	0.1296	−39.1*	0.0598	−71.9*
At 100	0.1208	0.0959	−20.6*	0.1223	+1.2	0.0701	−41.9*	0.0310	−74.3*
At 200	0.0831	0.0626	−24.7*	0.0841	+1.3	0.0486	−41.5*	0.0205	−75.4*
At 500	0.0483	0.0333	−31.1*	0.0478	−1.0	0.0275	−43.0*	0.0116	−76.0*
At 1000	0.0305	0.0197	−35.5*	0.0276	−9.5*	0.0167	−45.1*	0.0071	−76.8*
Average precision (non-interpolated)									
	0.1086	0.0827	−23.85*	0.1013	−6.73*	0.040	−63.05*	0.020	−81.24*

* Indicates t-test significance at $p < 0.05$.

the cluster with the lowest AP also the cluster with the highest spread. Thus cluster diversity is relatively evenly spread between the higher and lower precision clusters under a de-RR partitioning scheme.

For an additional comparison, we did one more experiment in which the top 2000 documents from the fused list (the same 2000 that were used to create the k-means clusters) were split evenly into top/bottom halves. The documents that were ranked 1–1000 became the first thousand documents in one, obviously higher precision, cluster. The documents that were ranked 1001–2000 became the 1st to 1000th documents in a second cluster. And while the precision from the "better" cluster is (by definition) equivalent to the precision from the initial list from which it was created, we also found that the spread of the bottom, "worse" cluster was larger only 81 times out of 150 (54%, sign test $p > 0.10$). Thus it is not the case that the second search collaborator can simply start further down in the ranked list to get a greater amount of diversity than the first search collaborator. More intelligent algorithmic mediation is necessary to support the team's activities.

These results indicate that while it is possible to create naive partitions from collaboratively created (fused) result sets, neither de-RR splitting nor the top 1000/bottom 1000 splitting will satisfy both Gatherer and Surveyor. De-RR comes close to creating a cluster that fulfills the Surveyor role, but from Table 3 we see that it cannot fulfill the Gatherer role. Top/bottom splitting, on the other hand, can fulfill the Gatherer role, in that one of the clusters is guaranteed to be equivalent to the initial fused list (which has significantly higher AP than either original list). However, it cannot fulfill the Surveyor role.

Our merge-split approach, on the other hand, creates partitions that support both Gatherer (high precision) and Surveyor (high diversity) roles. While diversity, especially in the absence of relevance, is not a typical information retrieval evaluation metric, we argue that this approach lets the search team more readily discover important clues that will guide further search iterations. For example, a more diverse set of results can help a team discover where their search may be failing by being able to see a broader range of information that is being retrieved by the system. The goal is to support the discovery of the kinds of "gaps in knowledge" that Marchionini describes (Marchionini, 2006). By looking at a more diverse result set, the Surveyor may identify terms to avoid in future searches, or discover potentially useful documents or search terms. By allowing one team member to focus primarily on that task, rather than mixing different goals, we believe that the team as a whole can work more effectively. Future work, though, will have to address the effect of that diversity on the iterative, interactive nature of the entire search session. We have, at least, shown that with two searchers there is merit to the role-based results redistribution. Overall, one need not sacrifice precision for diversity, as one of the users will still be given access to the precision-optimized cluster.

On the surface, these results appear to be at odds with those reported by Joho et al. (2009). There are, however, important differences between the two approaches that can account for the apparent discrepancy. Joho et al. clustered 300 documents to simulate assignments to multiple collaborating team members, and did not observe a difference in recall between round-robin and cluster-based assignment of documents to simulated searchers. Their methodology, however, makes that result unsurprising because they appear to be re-pooling results from all k team members after n documents per team member have been examined. Since they are measuring recall at $n * k$ documents, and the total number of relevant documents does not change, one would expect little difference in recombined team pools no matter if the initial assignment was clustered or round-robin – the team as a whole will get through most if not all of the same best-scoring documents in both cases. In our case, we are considering precision and diversity. The asymmetry of roles (a consideration that is absent in their analysis) makes it unnecessary to recombine the results assigned to each team member because the goals are ultimately different.

5. Future work

This paper presented an initial investigation into algorithmic support for asymmetric roles in collaborative exploratory search. We postulated a collaboration between two people, and modeled one cycle of the collaboration. From this point, several directions for extension are possible. For instance, our simple clustering approach could be enhanced to support more than one Gatherer. This can be achieved by employing a Scatter–Gather-style mechanism (Hearst & Pedersen, 1996) to identify more than one high-precision cluster. Other means of determining semantic facets could also be used to determine which documents are shown to whom. In particular, this approach may be useful for mediating collaboration among experts in different domains.

Another aspect that needs to be investigated is how to select which of the collaborating users' queries should be merged. A number of strategies are possible, including the most recent, the ones with the deepest judgments of relevance, or combinations of the two using a scheme similar to that described by Pickens et al. (2008) that combines relevance with recency.

Currently there is no mechanism for assigning a specific role to a specific person. One person will arbitrarily receive the Gatherer results, the other the Surveyor. We do not believe that this hinders the collaborators' ability to carry out a successful information seeking session. Furthermore, evidence in this paper leads us to believe that both the users and the system should be able to tell the result sets apart after examining a dozen results from each set (after a few documents have been marked relevant). If users wish to swap lists (roles) after at that point, the interface should allow them to. However, we acknowledge that there is a limit to simulation of collaborative information seeking environments and plan to test these ideas in future users studies.

It should be possible to allow users to request this computation manually by submitting specific queries to be processed in the manner described in this paper. These techniques could also be applied to the activity of a single user who "collaborates with himself" by combining queries from his search history. In this case, the two clusters can be labeled as different kinds of suggestions. For example, the label 'focused' can be applied to the set similar to the user's query with the most (positive) relevance judgments, and the label 'diverse' can be applied to the other cluster.

6. Conclusions

In this paper, we described one round, one stage, of an algorithmically-mediated role-based collaborative search process. Further work is needed to bring this into a full interactive session, but our preliminary results are encouraging. With our experiments on TREC collections, we demonstrated an effective mechanism to redistribute the results for Gatherer and Surveyor roles in a collaborative IR situation. We showed that while naive approaches to partitioning data may satisfy one or the other role's requirements, our proposed mechanism satisfies the needs of both roles.

Taking a step back, we wish to point out the importance of having such different roles. While in some cases collaboration with identical roles is appropriate, asymmetric roles can be effective in modeling certain kinds of collaboration. Both team members in the same role may get the benefit of division of labor in a collaborative setup, but it is the ability to combine diverse actions and inputs that gives collaboration its real strength (Shah et al., 2009). In other words, it is through such a collaboration with asymmetric roles and responsibilities that we can hope to make the *whole greater than the sum of the parts.*

An important assumption of our work is that people with the same information need tend to express similar queries, eliminating the possibility of completely different queries for a given topic. This assumption is a driving force for several of the works in collaborative search, such as Smyth, Balfe, Briggs, Coyle, and Freyne (2003). If, however, the queries are very different with little or no overlap in their result sets, the merge aspect of the algorithm technique may boil down to round robin merging of the two lists, but the splitting aspect will be unaffected. The Gatherer role may not be more precise than either person searching alone, but the Surveyor will still be optimized for diversity. As a degenerate example, the algorithm is capable of starting with a single ranked list created by one user's query and split it into Gatherer and Surveyor clusters.

In the work reported here our focus was on creating a mechanism that can enable combining and redistributing the search results to meet the criterion of asymmetric user-roles in a collaborative IR scenario. We discussed a situation of Gatherer–Surveyor and showed how we could facilitate the above requirement with algorithmic mediation. It is possible to define such user roles in other ways. For instance, we could divide the results according to the facets that they cover and present each user with a different facet. Such results can be obtained by performing aspectual or semantic clustering instead of the kind of clustering that we have proposed here. Imagine for a query such as "elephants", one cluster might be "African", the other "Asian". We can then have one user explore one path, the other user explore the other. This is just one of many possible ways to structure the collaborative interaction. Given a situation, one can decide to choose the kind of merging and splitting algorithms that would work the best for that context. We hope this work has shown a promising direction of reaping the benefits of role-based collaborative IR.

References

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734–749.

Aslam, J. A., & Montague, M. (2001). Models for metasearch. In: *Proceedings of ACM SIGIR* (pp. 276–284).
Cabri, G., Leonardi, L., & Zambonelli, F. (1999). Supporting cooperative WWW browsing: A proxy-based approach. Technical report, University of Modena, Italy.
Denning, P. J. (2007). Mastering the mess. *Communications of the ACM, 50*(4), 21–25.
Denning, P. J., & Yaholkovsky, P. (2008). Getting to we. *Communications of the ACM, 51*(4), 19–24.
Donath, J. S., & Robertson, N. The sociable web. In *Proceedings of WWW conference, CERN*, Geneva, Switzerland, May 25–27.
Foley, C. (2008). Division of labour and sharing of knowledge for synchronous collaborative information retrieval. PhD thesis, Dublin City University, June 2008.
Freyne, J., Smyth, B., Coyle, M., Balfe, E., & Briggs, P. (2004). Further experiments on collaborative ranking in community-based web search. *Artificial Intelligence Review, 21*, 229–252.
Fu, X., Kelly, D., & Shah, C. (2007). Using collaborative queries to improve retrieval for difficult topics. In *Proceedings of ACM SIGIR*, Amsterdam, The Netherlands, July 23–27 (pp. 879–880).
Gerosa, L., Giordani, A., Ronchetti, M., Soller, A., & Stevens, R. (2004). Symmetric synchronous collaborative navigation. In *Proceedings of the 2004 IADIS international WWW/Internet conference*, Madrid, Spain, October 6–9.
Golovchinsky, G., Pickens, J., & Back, M. A taxonomy of collaboration in online information seeking. CoRR abs/0908.0704. <http://arxiv.org/abs/0908.0704>.
Golovchinsky, G., Qvarfordt, P., & Pickens, J. (2009). Collaborative information seeking. *IEEE Computer, 42*(3).
Graham, T.C.N. (1997). GroupScape: Integrating synchronous groupware and the World Wide Web. In *Proceedings of INTERACT* July 14–18. Sydney, Australia: Chapman and Hall.
Greenberg, S., & Roseman, M. (1996). GroupWeb: A WWW browser as real time groupware. In *Proceedings of SIGCHI conference on human factors in computing systems*, November 16–20 (pp. 271–272). Boston, MA: ACM Press.
Hansen, P., & Jarvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. *Information Processing and Management, 41*, 1101–1119.
Hearst, M., & Pedersen, J. (1996). Reexamining the cluster hypothesis: Scatter/gather on retrieval results. In *Proceedings of ACM SIGIR*, Zurich, Switzerland (pp. 76–84).
Hyldegárd, J. (2006). Collaborative information behaviour-exploring Kuhlthau's information search process model in a group-based educational setting. *Information Processing and Management, 42*, 276–298.
Joho, H., Hannah, D., & Jose, J. M. (2009). Revisiting IR techniques for collaborative search strategies. In *Proceedings of ECIR*, Toulouse, France, April 6–9.
Kautz, H., Selman, B., & Shah, M. (1997). Referral web: Combining social networks and collaborative filtering. *Communications of the ACM, 40*(3), 63–65.
Keller, R. M., Wolfe, S. R., Chen, J. R., Rabinowitz, J. L., & Mathe, N. (1997). A bookmarking service for organizing and sharing URLs. *Computer Networks and ISDN Systems, 29*, 1103–1114.
Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science and Technology, 42*(5), 361–371.
Kuhlthau, C. C. (2005). Towards collaboration between information seeking and information retrieval. *Information Research, 10*(2).
Laurillau, Y., Nigay, L. (2002). CoVitesse: A groupware interface for collaborative navigation on the WWW. In *Proceedings of the ACM conference on computer supported cooperative work (CSCW)*, New Orleans, LA, USA, November 16–20 (pp. 236–240).
Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing, 7*(1), 76–80.
Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM, 49*(4), 41–46.
Menkov, V., Neu, D. J., & Shi, Q. (2001). AntWord: A collaborative web search tool. In *Workshop on distributed communities on the web (DCW)*, Quebec City, Quebec, Canada, June 19–21.
Morris, M. R. (2008). A survey of collaborative web search practices. In *Proceedings of SIGCHI conference on human factors in computing systems*, Florence, Italy, April 5–10.
Morris, M. R., & Horvitz, E. (2007). SearchTogether: An interface for collaborative web search. In UIST, Newport, RI, October 7–10 (pp. 3–12).
Morris, M. R., Teevan, J., & Bush, S. (2008). Enhancing collaborative web search with personalization: Groupization, smart splitting, and group hit-highlighting. In *Proceedings of CSCW 2008* (pp. 481–484).
Paul, S., & Morris, M. R. (2009). CoSense: Enhancing sensemaking for collaborative web search. In *Proceedings of CHI 2009* (pp. 1771–1780).
Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., & Back, M. (2008). Algorithmic mediation for collaborative exploratory search. In *Proceedings of ACM SIGIR*, Singapore, July 20–24.
Root, R. W. (1988). Design of a multi-media vehicle for social browsing. In *Proceedings of the conference on computer-supported cooperative work (CSCW)* (pp. 25–38).
Shah, C., Marchionini, G., & Kelly, D. (2009). Learning design principles for a collaborative information seeking system. In *The proceedings of CHI 2009*, April 4–9. Boston, MA.
Shah, C. (2009). Lessons and challenges for collaborative information seeking (CIS) systems developers. In *The proceedings of collaborative information behavior workshop at GROUP 2009*, May 10, Sanibel Island, Florida.
Shaw, J. A., & Fox, E. A. (1994). Combination of multiple searches. In *Text (REtrieval) conference* (pp. 105–108).
Smeaton, A. F., Lee, H., Foley, C., McGivney, S., & Gurrin, C. (2006). Físchlár-DiamondTouch: Collaborative video searching on a table. In *Proceedings of SPIE* (Vol. 6073). <http://dx.doi.org/10.1117/12.643077>.
Smyth, B., Balfe, E., Briggs, P., Coyle, M., Freyne, J. (2003). Collaborative web search. In *Proceedings of the international joint conference on artificial intelligence (IJCAI)* August 9–15 (pp. 1417–1419). Morgan Kaufmann: Acapulco, Mexico.
Smyth, B., Balfe, E., Briggs, P., Coyle, M., & Freyne, J. (2003). Collaborative web search. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 1417–1419), August 9–15. Acapulco, Mexico.
Twidale, M. B., & Nichols, D. M. (1996). Collaborative browsing and visualisation of the search process. In *Proceedings of Aslib* (Vol. 48, pp. 177–182).
Twidale, M. B. T., Nichols, D. M. N., & Paice, C. D. (1997). Browsing is a collaborative process. *Information Processing and Management, 33*(6), 761–783.
Wittenburg, K., Das, D., Hill, W., & Stead, L. (1995). Group asynchronous browsing on the World Wide Web. In *Proceedings of WWW conference*, Boston, MA, December 11–14.