

Cerchiamo: a collaborative exploratory search tool

Gene Golovchinsky, John Adcock, Jeremy Pickens, Pernilla Qvarfordt, Maribeth Back

FX Palo Alto Laboratory, Inc.
3400 Hillview Ave, Bldg. 4
Palo Alto, CA 94304
{gene, adcock, pickens, pernilla, back}@fxpal.com

ABSTRACT

We describe Cerchiamo, a collaborative exploratory search system that allows teams of searchers to explore document collections synchronously. Working with Cerchiamo, team members use independent interfaces to run queries, browse results, and make relevance judgments. The system mediates the team members' search activity by passing and reordering search results and suggested query terms based on the teams' actions. The combination of synchronous influence with independent interaction allows team members to be more effective and efficient in performing search tasks.

Author Keywords

Collaborative search, information retrieval, exploratory search, CSCW, video search, interactive search, search algorithms

ACM Classification Keywords

H.3.3 Information Search and Retrieval

INTRODUCTION

The term “collaboration” or “collaborative” has been applied to systems that allow multiple people to interact in some ways. In the area of information retrieval for instance, recommendation systems where information is retrieved based on browsing or querying behaviors similar to those of prior users are often labeled as “collaborative.” We believe that whereas cooperation or correlation apply to a broader set of computer-mediated activities, “collaboration” implies explicitly shared intent. We built Cerchiamo (“we search” in Italian), a collaborative exploratory search system, that allows teams of searchers to coordinate their search activity with respect to a common goal [5]. The system lets each person use a dedicated search interface independently of other team members, but routes information derived from each team member to others in appropriate ways.

In this paper, we review some relevant work, describe our role-based approach to information seeking, and then describe the key components of Cerchiamo.

RELATED WORK

We created a taxonomy of collaboration in information seeking [2] that categorizes systems based on user intent (implicit vs. explicit), synchrony, depth of collaborative mediation and location. Thus recommender systems (e.g., [3]) are classified as asynchronous, implicit, with deeply

algorithmic mediation and non-co-located. Collaborative search systems such as SearchTogether [4] and Físchlár-DiamondTouch [6] are synchronous, explicit in intent, with interface-only mediation and co-located users. Cerchiamo differs from both because it is synchronous, explicit and with deeply algorithmic mediation. It is agnostic with respect to co-location. The distinction in the mediation dimension is important for two reasons: it frees team members to pursue complementary strategies, and yet allows them to influence and inform each others' interactions.

In our algorithmically-mediated system both users, acting in each of their prescribed roles, influence each other in real-time. This influence is seamless, integrated and bi-directional. Neither user must actively disengage from his or her current search activities in order to share information with the search partner. The underlying search algorithm takes into account the search activities of both users and redistributes information and adjusts underlying retrieval algorithm parameters based on both users' roles.

ROLES

Many kinds of roles are relevant to activities around information seeking, including peers, domain-expert vs. search expert, and Prospector vs. Miner. We chose to focus on the latter set. The Prospector discovers potentially promising directions of exploration, whereas the Miner digs deeper into the results to find information that may not be initially obvious. Our evaluation [5] of these roles in the context of the TRECVID interactive competition suggests that the combination is more effective and efficient at finding relevant documents (particularly when search topics are sparse) than post-hoc merging of search results generated by two searchers working independently.

CERCHIAMO PROTOTYPE SYSTEM

The system architecture consists of three parts: the User Interface Layer, the Regulator Layer, and the Algorithmic Layer (Figure 1). System components communicate through a web service API, and can be combined in different ways: the single shared display in a co-located setting can be replaced by separate displays in remote locations, showing the same information. In our demonstration searchers are co-located.

User Interface Layer

Cerchiamo contains three user interfaces: MediaMagic [1] for the Prospector to issue queries and browse results, a rapid serial visualization result browsing user interface (RSVP) for the Miner to make relevance judgments, and a

shared display containing information relevant to the progress of the search session.

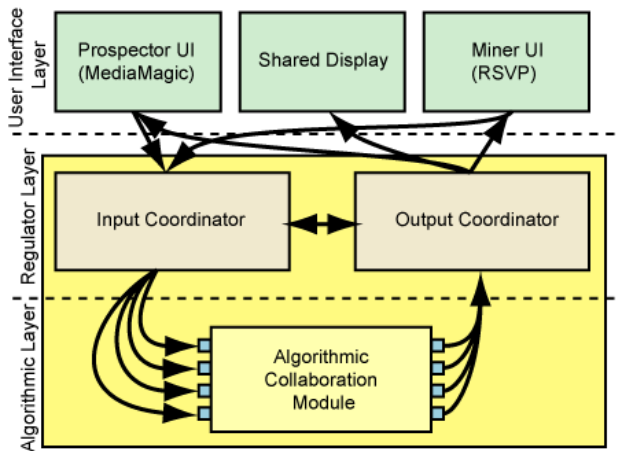


Figure 1. Collaboration architecture

Regulator Layer

The regulator layer is responsible for managing influence. Influence from the Prospector to the Miner flows in the form of retrieved, but unseen (un-judged) documents. The Prospector issues new queries and makes relevance judgments throughout a search session. Rather than examining all retrieved documents, the Prospector may decide to issue a new query to cover as much of the search space as possible. At this point, the mediation algorithms incorporate the Prospector's retrieved but unseen documents into the list presented to the Miner. The order of documents presented to the Miner is different from the order presented to Prospector. As the Prospector continues issuing new queries and making more relevance judgments, the Miner's queue gets continually reshaped and reordered.

Influence also flows from the Miner to the Prospector in the form of query term suggestions. The system identifies terms characteristic of documents that were judged relevant by the Miner, and makes them available to the Prospector for reuse in future queries.

The difference between this collaborative system and typical query term suggestion is that in latter systems, once a Prospector has moved on from the current query pathway, no new documents from that pathway are seen, and therefore the relevance weights on terms extracted from that pathway do not change. In our system, the Miner continues to explore the most promising pathways, and therefore constantly updates the list of query term suggestions based on newly-found relevant documents. Furthermore, the Miner does not have to actively suggest these terms to the Prospector; the suggestion happens automatically and in real-time, via algorithmic mediation [5]. In this manner the Prospector is able to both be aware of the Miner's current activity as well as see term suggestions that the Prospector would not have seen had he been working alone.

Algorithmic Layer

The algorithmic layer consists of a number of functions for combining searchers' activities to produce documents, rankings, query suggestions, and other information relevant to the search. It performs basic searches, and generates raw search results, transformed search results based on input from multiple users, and query terms that characterize the current state of the collaboration. The algorithms are described in detail in [5].

CONCLUSION

We present Cerchiamo, a collaborative exploratory search system, which allows users to explicitly and synchronously collaborate while being supported by deeply algorithmic mediation. In comparison to other synchronous collaborative search systems, Cerchiamo allows the users to pursue the search task at their own pace without needing to synchronize each search activity with their collaborators. Future research and refinement in algorithms, role-based system design, and interface design should further improve the prospect of collaborative search. Cerchiamo illustrates the need to build search system with theories and practices from the fields of both HCI and IR. Collaborative search systems need to support the collaborative user experience as well as algorithms mediating the searchers' activities.

REFERENCES

1. Girgensohn, A., Adcock, J, Cooper, M, and Wilcox, L. A Synergistic Approach to Efficient Interactive Video Retrieval, In *Proceedings of INTERACT 2005*, LNCS 3585, (September 2005) pp. 781-794
2. Golovchinsky, G., Pickens, J., and Back, M. A Taxonomy of Collaboration in Online Information Seeking. In *Proceedings of the 1st International Workshop on Collaborative Information Retrieval*. June 20, 2008. Available online at <http://www.fxpal.com/?p=abstract&abstractID=454>
3. Linden, G., Smith, B. and York, J. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 7 (1). 76-80.
4. Morris, M. R. and Horvitz, E. 2007. SearchTogether: an interface for collaborative web search. In *Proceedings of UIST '07*. ACM, New York, NY, 3-12.
5. Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., and Back, M. Algorithmic Mediation for Collaborative Exploratory Search. In *Proceedings of SIGIR 2008*, July 22-25, 2008, ACM Press.
6. Smeaton, A.F., Lee, H., Foley, C., McGivney, S. and Gurrin, C., Físchlár-DiamondTouch: Collaborative Video Searching on a Table. In *Proceedings of the SPIE Electronic Imaging-Multimedia Content Analysis, Management, and Retrieval* (2006).
7. Smyth, B., Balfe, E., Freyne, J., Briggs, P., Coyle, M. and Boydell, O. Exploiting Query Repetition and Regularity in an Adaptive Community-Based Web Search Engine. *User Modeling and User-Adapted Interaction*, 14 (5). 383-423..