

Collective Information Seeking: Supporting Search with Low-Cost Activity Sharing and Collaboration

Scott Bateman, Carl Gutwin, Gordon McCalla
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
<firstname.lastname>@usask.ca

Ryen W. White
Microsoft Research
Redmond, Washington, USA
ryenw@microsoft.com

ABSTRACT

When searching for information, people often seek help from others. However, while people can benefit from communicating with others, they can usually satisfy their information needs, to some degree, without help. Because establishing explicit collaborations is often seen as onerous, there are many missed opportunities where shared experiences could save time and effort. In this work, we propose *Collective Information Seeking*, which integrates theory of group work with previous research in Social Navigation and Collaborative Information Seeking. We propose that by sharing and visualizing activities within groups of information seekers a low-cost form of collaboration is made possible. Shared activities provide both a form of implicit communication that can provide guidance, and provide cues and context facilitating entry into direct collaboration at a later time.

Categories and Subject Descriptors

H.5.3 Group and Organization Interfaces: CSCW.

General Terms

Design, Human Factors, Theory.

Keywords

Collective information seeking, social navigation, collaborative information seeking.

1. INTRODUCTION

Current information seeking tools have primarily been designed for single users working alone [11]. However, even with single-user tools, information seekers explicitly communicate with others: before searching, to help describe and focus what is being looked for; during search, to receive advice on what avenues to pursue; and after search, to share what has been found [2].

Collaboration in many information seeking situations begins with an *informal* form of collaboration [1]. Informal collaborations are characterized by being *opportunistic* and *discretionary* [8]. Opportunistic collaborations arise as opportunities present themselves and are grounded in being aware about the activities of other people and what they might be able to provide guidance on. Discretionary collaborations are those that do not need to occur at all, as the task can be completed, to some extent, individually.

Consider the situation where a user, John, wants to find a hotel in a safe neighborhood in Vancouver. He is unaware of the fact that several of his colleagues have recently travelled to Vancouver and had good experiences with their accommodations. John's colleagues are not around for him to quickly ask if they had any relevant Vancouver travel advice to share. John wants to get "booking a hotel" off his to-do list, so he picks the first one he finds that seems good enough. In this example, John has decided

the overhead of emailing or tracking down his colleagues, on the off chance they might have something to recommend, is too high. Situations such as this lead to missed opportunities and lead users to be less efficient. In such situations users could benefit from the knowledge and guidance of others, but without the right tools, acquiring and sharing information is often too costly.

In this position paper we highlight that for guidance and experiences to be widely shared between individual information seekers, systems must be designed that take into account how communication and collaboration naturally occur in group work. In particular, most communication in groups is informal; it is both unplanned and discretionary in nature. In informal collaborations, people opportunistically make use of cues in work environments, such as the documents someone has on their desk or overheard conversations, to provide an awareness of their past and current activities. Being aware of the activities others are undertaking is an implicit form of collaboration, because it can inform decision making. However, being aware of what others have done can also be used as means to enter into more direct forms of explicit collaboration [10].

This paper proposes the design of *collective information seeking* (CIS) systems that provide opportunities to gather detailed information on the activities of other members in a group of familiar users. The benefits to having an appropriate awareness of others' activities are threefold.

- First, having access to others' activities (e.g., the searches people have recently performed) allow users to update their knowledge about what other people are currently doing and can provide advice on. For example, had a system informed John that he had a colleague who recently issued a "Vancouver hotel" query he would have known they might have advice.
- Second, being aware of the activities of others may allow users to collaborate implicitly, relieving the need to enter into a higher-cost explicit collaboration. For example, a system could display the trail John's colleague followed after issuing his "Vancouver hotel" query. John sees that one of his colleagues had last clicked on a link for the Pan Pacific Hotel and then went to the room booking page. John decides to book the hotel, knowing that this particular colleague has a good attention to detail and that this hotel will likely meet his needs too.
- Third, an up-to-date and contextual awareness of others' activities can provide cues and opportunities to enter into direct collaboration. For example, John may be working on other tasks, but a system updates him that another of his colleagues, Mary, is currently uploading pictures to her "Vancouver 2011" photo album. Learning that Mary has recently been to Vancouver and that she is not currently busy with work, he requests a browser sharing session with Mary. Mary agrees and

guides John through her photos of Vancouver. She points out a few good hotels that are close to landmarks in her pictures.

In these situations, because a system appropriately provided John with the details of his colleagues' activities, he was able to discover that they have information that could help him. Further, he was able to appropriately engage Mary in a low-cost communication to get a response to his question. Finally, in one situation John realized that the awareness afforded by the system was sufficient to help him make an informed decision on a hotel. In these examples, had the cost of getting details on other peoples' experiences or entering into communication been too high, John would have likely chosen to continue without help or abandon his information seeking task rather than engage with others. Because people view the advice needed to make decisions in information seeking tasks as secondary to the task itself, they are only willing to seek advice if it is quick and easy.

In the remainder of this paper we first introduce the concepts of *awareness* and *informal collaboration*, which help describe the gaps in the design space that CIS addresses. We then briefly characterize work from social navigation and collaborative information seeking, upon which this work builds. Throughout, we illustrate situations where a CIS system enables new forms of communication and collaboration in information seeking. Finally, we outline next steps for designing and building a CIS system as a browser extension for use on the open Web, and discuss the key issues that must be more fully understood: (i) Do people have enough overlap in their tasks and activities to provide opportunities for collaboration?; (ii) Do people find the activities of others useful?; (iii) What kinds of activities are most useful?; (iv) Are people willing to share their activities? (v) How can activities best be shared and displayed in the browser?

2. RELATED WORK

2.1 Awareness

“Awareness involves knowing who is ‘around,’ what activities are occurring, who is talking with whom; it provides a view of one another in the daily work environments. Awareness may lead to informal interactions, spontaneous connections, and the development of shared cultures... ([4], p. 541).”

Dourish and Bellotti described awareness as “an understanding of the activities of others, which provides a context for your own activity ([3], p. 107).” This means awareness information can help users refine their own tasks. For example, two collaborators who are scouring a website for hidden information can be aware of each other's activities, thus avoiding duplicating effort.

Awareness allows people to answer questions such as: What has been done here before? Who did it? And, when was it done? [7] Having these questions answered can relieve the need for higher-cost direct communications, but still allows users to receive helpful information. We consider this a form of implicit collaboration, where people make use of consequential communication from the activities of others to gain valuable information about their work and environment.

A recent study showed that duplication of effort was substantially reduced and that explicit communication between collaborators was more efficient without increasing cognitive load, when an appropriate awareness of collaborators actions was provided [12].

2.2 Informal Collaboration

Kraut and colleagues [10] describe informal interactions as being those that are not scheduled, and include communications that occur in an unplanned or completely spontaneous manner. Informal collaboration also includes collaborations that arise out of encountering cues in the environment, such as seeing someone and asking what they are working on. Gutwin et al. [8] highlight seven characteristics of informal collaboration that must be taken into account when designing collaborative systems.

Informal interactions are common in collocated groups. In a large scale categorization of communications it was found that 88% of conversations were informal [10]. Informal collaborations were also found to be the most frequent in a study of information seeking, where 62.9% collaborative, collocated searches in the workplace were unplanned [1].

Informal interaction is grounded in awareness of the work environment. People maintain awareness through the people, objects, and activities that their colleagues are undertaking in the environment. Informal collaboration is based upon having an up-to-date and rich awareness of other people and the environments they are working in [8].

Informal collaboration can be triggered by people, objects, actions, or interaction. Cues for informal collaboration include seeing a person or an object in the work environment, seeing someone doing a particular action, or overhearing a conversation.

Informal collaboration is often discretionary. Informal communication allows participants to make choices about whether they need to engage in them, and whether they ever need to occur at all. The communication may never need to happen at all, this allows individuals to decide if the cost of entering into communication is too high.

Informal interactions are easy to initiate. Due to informal communication being discretionary, engaging in collaboration must be low-cost or people will choose simply not to participate.

Entry into collaboration is negotiated. Beginning explicit collaboration requires multiple steps, including: seeing if the other party is available, and; determining if they are willing and able.

Collaborations often require interaction with task artifacts. Many tasks involve objects that need to be referred to and manipulated during communications [8]. Having an awareness of the past activities of others in working with task artifacts can facilitate collaboration [12].

2.3 Collaborative Information Seeking

Work in collaborative information seeking focuses on collaboration between users who have *explicitly shared information needs* [6]. This work has proposed a number of specialized systems for users to satisfy a shared need. Tools for facilitating collaboration under different situational constraints have been a focus. For example, shared tabletop displays have been proposed for collocated situations, or systems that organize collaborators' search trails for work taking place at different times and places (see [11] for a review).

Because this work has primarily focused on facilitating explicit collaborations, it proposes solutions that are geared towards more costly forms of communication, and not towards informal types of collaboration. Further, the best way to provide opportunities for entry into explicit collaborations has not been explored.

2.4 Social Navigation

Social navigation is a concept coined by Dourish and Chalmers [5] that characterizes an approach for navigation in information spaces. Social navigation is concerned with guiding people through a larger information space based on the interpreted actions, expressed thoughts or explicit recommendations of others. Dourish has said that the “navigation” in social navigation is best understood as an information seeking task [5], where information about community activity enables users to make better decisions.

Hill, et al. [9] originated the idea of “[capturing] on computational objects (e.g. documents, menus, spreadsheets, images, email) the events that comprise their use ([9], p. 3).” Their goal was to graphically depict these histories, giving the viewer a rich understanding of an object’s use. They referred to this as “computational wear”, alluding to the metaphor of the wear that objects accrue in the physical world. They developed read wear and edit wear tools that visually depicted the reading and editing history of all users working on a document. Their goal was to make apparent from glancing at the graphical traces of history answers to questions relating to awareness, such as: Who wrote that part of the document? Did my colleague read that part? What is the relative age of that paragraph? [9].

Work on computational wear led to social navigation systems that employ *social cues*, which visualize traces of activity of a community of users. The work of Wexelblat and Maes, called “Footprints”, annotated paths through websites by making popular paths more visually salient than others [13]. Such social cues offer a low-cost solution for an implicit form of collaboration. It is low-cost because usage data can be collected and presented without explicit actions by users. A study of the Footprints system showed users’ decisions were better informed; they were able to find target resources in significantly fewer steps.

In the Footprints work, and most work employing social cues that followed, individual activity traces are aggregated for all users, providing only the indicators for an entire community. This limits the information that can be drawn by the viewer. For example, had John only been able to see aggregated cues while looking for a hotel, he would not be able to distinguish the actions of his colleagues from those of other users. This differs from the original vision of Hill et al. who described a richer form of awareness that includes, for example, the identities of history creators.

3. DESIGN REQUIREMENTS

In addition to meeting the previously described requirements of informal collaboration, CIS is based on two core tenets.

First, the depiction of past activities must provide users with sufficient information that allow them to understand what and when particular activities were carried out, and who did them. The goal is to provide the viewer with sufficient information to enable them to discern whether or not an activity is of interest or not.

Second, tools must provide the appropriate form of collaboration given an opportunity. We believe that in many situations the visual cues of the past activities will be of interest to users and provide sufficient information to support tasks or otherwise provide awareness. However, situations will occur in which a user determines that someone else has information that could be helpful, but where the activities themselves do not carry sufficient information to allow an informed decision. Such situations require a more direct form of collaboration – which may be synchronous,

asynchronous, or both – to be undertaken in order to allow users to receive sufficient guidance to complete their tasks.

3.1 Displaying Activities

In this paper we have suggested that traces of activities be shared and displayed visually during information seeking tasks. We see two main ways that activities can be shared in web browsers: contextually and externally.

Contextually shared activity traces include those that are integrated as closely as possible to where the original interactions took place. For example, the edit and read wear tools visualize activities in the scrollbar, providing both the detail and position of the activities. Contextually integrating cues would allow users to encounter and make use of shared activities with little effort. For example, annotations on hyperlinks could indicate that the link has been selected by several acquaintances. Mousing over the annotation could provide details of who used the hyperlink and when. One could even imagine users employing contextual cues for new forms of communication. For example, John and his friends use a system that visualizes the visits to hyperlinks using “wear”. John wants his friends to notice his updated photo album on his website, so he repeats the path to his photo album several times so that it will stand out to his friends.

Contextually shared activities suffer from the disadvantage that users must share a resource (such as a webpage) in order for them to be displayed. Because of the vastness of the Web, these may be encountered infrequently within a group of familiar users. Further, in some of the previous examples John benefited from knowing what Mary was doing without sharing the same task. We believe that it will also be important to provide traces of activities that are external to a user’s current context. These could be displayed in a web history-style tool that would allow user to review or search recent activities, or a scrolling ticker could display the current activities of others. The issue of infrequently encountered traces may also be partially addressed by determining exactly when and how traces may be most usefully displayed. For example, activities could be aggregated and displayed for a domain rather than just for a URL. Such an approach would allow the traces of other users to be encountered more frequently.

3.2 Forms of Collaboration

CIS is centered on sharing activity traces both for implicit collaboration (using the traces themselves for guidance) and to enable entry into explicit collaboration (receiving guidance directly from another user). Thus far we have described asynchronous situations where traces have provided sufficient information to avoid collaboration and synchronous situations where traces (both historic and present) have provided an opportunity for easily transitioning into synchronous collaboration. Being able to transition to synchronous collaboration is important as it can enable users to receive and provide the most information, by communicating directly with others within the context of a task. However, since the focus of CIS is on providing opportunities for collaboration it must be recognized that in many situations that entry into synchronous and explicit forms of collaboration may not be possible. Recall the example when John, while looking for a hotel, was able to engage Mary who happened to be uploading her Vancouver photos. While this situation maybe plausible it is not always likely, because: 1) Mary needs to be working on a task (or has worked on a task) similar to John; 2) Mary needs to be online at the same

time as John; 3) Mary needs to be willing and able to collaborate with John at that particular time. These three requirements describe a rather narrow set of circumstances, and so it is most likely that collaborations will most likely occur in asynchronous situations; hence the focus on implicitly left traces.

However, there may also be situations where explicit forms of collaboration could be used in asynchronous scenarios. For instance, when Mary is not online at the same time as John or is not currently able to collaborate. In these situations, we could imagine asynchronous annotations that are left to seek or provide guidance when it is convenient for others to do so. Such annotations could be used to enable explicit asynchronous collaborations, enabling a fuller set of situations to be addressed.

4. DISCUSSION

In this position paper, we have argued for a different consideration of how collaboration in information seeking tasks can be supported. The basis of this view is that in order for familiar users to work together in information seeking tasks, systems must be designed that take into account how people engage in collaboration in many real-world situations. While collaborative information seeking systems have been proposed, they have been focused on situations where users have explicitly shared information needs and are engaged in a collaboration that tends towards formal. Because collaboration in information seeking is most often informal, systems must provide sufficiently low-cost solutions or users will not engage in collaboration. We believe that the appropriate collection and display of past and present activities between users provides a low-cost solution that would provide opportunities where collaboration can naturally arise.

While we believe that our vision of collective information seeking is feasible there are still several outstanding issues that need to be addressed and understood.

We are proposing the collection and display of traces of activities between familiar users. However, it is not known whether or not familiar users working in groups will have sufficient overlap in their tasks that sharing traces of activities will ever be useful for collaboration. This also raises the issue of whether or not users are willing to share their traces, and if so under what conditions will they be willing to share them. In particular there are important privacy issues that must be accounted for. For example, John may not want his colleagues to be able to see that he is browsing a competitor's employment page. We believe such concerns can be dealt with, but the exact methods employed have tradeoffs in terms of the amount of effort required and the coverage of activities that can be made available. For instance, users could explicitly select which traces they are willing to share and with whom to share them. However, such an approach would likely lead to very little sharing because of the effort involved. Other approaches such as whitelisting or blacklisting domains, and explicit browser modes could lead to wider spread sharing with reduced effort, but could lead to users mistakenly sharing information they would like to keep private. A CIS system user should always be able to delete traces that they decide should be private.

Further, given that opportunities for sharing activities exist, do people find them useful, and which types and form of activities are most useful? For example, one could imagine several granularities of activities being presented: low-level actions, could be presented such as increasing the visual saliency of links other users have

taken most often; or higher-level actions, such as a larger task of planning a trip. Finally, given a set of useful activities, exactly how they should be displayed is an open issue.

While there remain important open issues, we believe that they can be explored and dealt with. We are planning to design and build a CIS system iteratively within a real working group over an extended period of time. The system will be a web browser extension and will first support the most basic forms shared activities. For example, indicators of activities such as webpages visited could be displayed in a scrolling ticker, and direct collaboration could be achieved through sharing screen shots and providing basic telepointers. Such simple approaches will be rolled out to users, and successes and failures will be built upon iteratively. We believe this will allow us to refine our theory of collective information seeking and provide extended guidelines for the design of systems for information sharing and collaboration between information seekers.

5. ACKNOWLEDGMENTS

This research is partially funded by NSERC. The first author is partially supported by a Walter C. Sumner Memorial Fellowship.

6. REFERENCES

1. Amershi, S. & Morris, M.R. CoSearch: A system for collocated collaborative Web search. *Proc. CHI '08*, 1647-1656.
2. Chi, E.H. Information seeking can be social. *IEEE Computer*, 42(3), March, 2009, 42-46.
3. Dourish, P. & Bellotti, V. Awareness and coordination in shared workspaces. *Proc. CSCW '92*, 107-114.
4. Dourish, P. & Bly, S. Portholes: Supporting awareness in a distributed work group. *Proc. CHI '92*, 541-547.
5. Dourish, P. & Chalmers, M. Running out of space: Models of information navigation. *Proc. HCI '94*.
6. Golovchinsky, G., Qvarfordt, P. & Pickens, J. Collaborative information seeking. *IEEE Computer*, 42(3), 2009, 47-51.
7. Gutwin, C. & Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work*, (11)3, (Nov. 2002), 411-446.
8. Gutwin, C., Greenberg, S., Blum, R., Dyck, J., Tee, K. & McEwan, G. Supporting informal collaboration in shared-workspace groupware. *Journal of Universal Computer Science*, 14(9), 1411-1434.
9. Hill, W.C., Hollan, J.D., Wroblewski, D. & McCandless, T. Edit wear and read wear. *Proc CHI '92*, 3-9.
10. Kraut, R., Fish, R., Root, R. & Chalfonte, B. Informal communication in organizations: Form, function, and technology. In *Readings in Groupware and Computer Supported Cooperative Work*, Morgan Kaufmann, 1993.
11. Morris, M.R. & Teevan, J. *Collaborative Web Search: Who, What, Where, When, and Why*. Morgan and Claypool, 2010.
12. Shah, C. & Marchionini, G. Awareness in collaborative information seeking. *JASIST*, 61(10), 2010, 1970-1986.
13. Wexelblat, A. & Maes, P. Footprints: History-rich tools for information foraging. *Proc. CHI '99*, 270-277.