Collaborative Information Seeking: Art and Science of Achieving 1+1>2 in IR

Chirag Shah
School of Communication & Information
Rutgers University
4 Huntington St, New Brunswick, NJ 08873, USA
chirags@rutgers.edu

ABSTRACT

Traditional IR techniques, systems, and methods that assume an individual searcher are often shown to be inadequate for addressing search problems that are multi-faceted and/or too complex or difficult for individuals. The next big leap in information seeking/retrieval could happen by considering social and collaborative aspects of search. In this half-day tutorial, this concept, along with some of the foundational works and latest developments in the field of collaborative information seeking (CIS) will be presented. Specifically, the course will introduce the student to theories, methodologies, and tools that focus on information retrieval/seeking in collaboration. The student will have an opportunity to learn about the social aspect of IR with a focus on collaborative search or CIS situations, systems, and evaluation techniques. The three hours will be divided as: (1) introduction to group-based IR models, approaches, and systems; (2) back-end of CIS systems with system-focused mediation and front-end with user-focused mediation; and (3) evaluation of CIS systems/approaches, prediction and recommendations with collaborative aspects of IR, and future directions. The attendees will be given a course-pack that will include a reference list, an annotated bibliography of seminal works in the field, and depictions of relevant models/frameworks.

Length: Half day (3 hours)

Target Audience: Those interested in social and collaborative aspects of IR (from both academia and industry)

Prerequisite: A general understanding of web services, IR systems, and content analysis

Presenter: Chirag Shah

Affiliation: Rutgers University, USA

Address:
4 Huntington St, New Brunswick, NJ 08901
Phone: (848) 932-0887
Fax: (732) 932-6916
http://comminfo.rutgers.edu/ chirags
chirags@rutgers.edu

Bio:
Chirag Shah is an associate professor in both the School of Communication & Information (SC&I) and the Department of Computer Science at Rutgers University, USA. Shah received a PhD in Information Science from the University of North Carolina (UNC) at Chapel Hill. He holds an MTech, Computer Science & Engineering from IIT Madras, India and an MS, Computer Science from UMass Amherst. His research interests include various aspects of interactive information retrieval/seeking, especially in the context of online social networks and collaborations, resulting in two streams of works - Collaborative Information Seeking (CIS) and Social Information Seeking (SIS). This research is supported by grants from the US National Science Foundation (NSF), the Institute of Museum and Library Services (IMLS), Google, and Yahoo! At Rutgers, he directs the InfoSeeking Lab (http://infoseeking.org/). He has published two books on CIS - one as a solo author [9] and the other as a co-editor [5]. He was a guest editor for the IEEE Computer Special Issue on CIS published in March 2014.

SIGIR '16, July 17-21, 2016, Pisa, Italy

DOI: http://dx.doi.org/10.1145/2911451.2914801

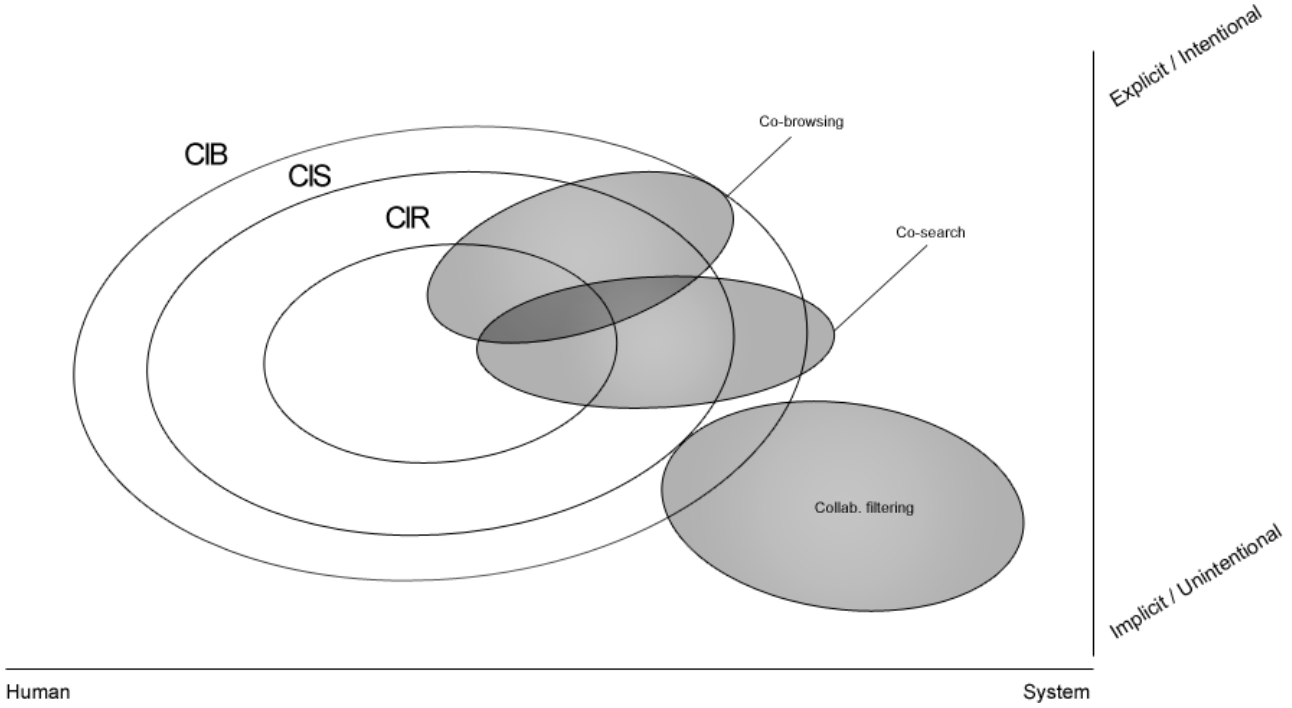

Figure 1: Depiction of CIS and related topics using the dimensions of human-system and explicit-implicit collaboration.

1. OVERVIEW AND RATIONALE

The course will introduce the student to theories, methodologies, and tools that focus on information retrieval/seeking in collaboration. The student will have an opportunity to learn about the social aspect of IR with a focus on collaborative search or collaborative information seeking (CIS) situations, systems, and evaluation techniques.

Traditionally, IR is considered an individual pursuit, and not surprisingly, the majority of tools, techniques, and models developed for addressing information need, retrieval, and usage have focused on single users. The assumption of information seekers being independent and IR problem being individual has been challenged often in the recent past. This course will introduce such works to the students, with an emphasis on understanding models and systems that support collaborative search or browsing. To put CIS in perspective, the course will show the students how various related concepts, such as collaborative information behavior (CIB), co-browsing, co-search, collaborative filtering, can be placed on the dimensions of human-system and explicitness-implicitness along with CIS and CIR (Figure 1) for exploration and developmental needs, as well as evaluation purposes.

Specifically, the course will (1) outline the research and latest developments in the field of collaborative IR, (2) list the challenges for designing and evaluating collaborative IR systems, and (3) show how traditional single user IR models and systems could be mapped to those for CIS. This will be achieved through introduction to appropriate literature, algorithms and interfaces that facilitate CIS, and methodologies for studying and evaluating them. Thus, the course will offer a balance between theoretical and practical elements of CIS.

2. COURSE OBJECTIVES

The course will have the following broad objectives:

1. Define collaborative information seeking (CIS) and differentiate it with related terms.
2. Identify situations and motivations for CIS.
3. List various dimensions of collaborative work/systems/method.
4. Discuss system-mediated and user-mediated CIS with examples.
5. Enumerate essential and desired features of CIS systems.
6. Outline an evaluation framework for CIS.

3. RELEVANCE TO THE IR COMMUNITY

With an increasing number of researchers from the fields of IR, HCI, and social media coming together, there has been an unprecedented interest to CIS/CIR and related issues in the recent years. While the concept of people working together in solving IR problems has been explored before, it has made its strong comeback in the past few years. This area has gained significant momentum through a number of workshops and other professional events, as well as special issues of journals. And yet, we have just scratched the surface. With the advent of social media and collaboratively generated content (CGC) everywhere, it is time the IR community takes a serious look at this topic that is increasingly achieving more relevance and importance. With its applicability to numerous domains including education, legal, health, and business, studying various issues relating to CIS/CIR is very timely and relevant to the IR community.

4. OUTLINE OF THE TUTORIAL

The three-hour course will follow a general structure given below.

- Hour-1: Discuss situations for CIS; Introduce a few group-based information seeking models and systems (e.g., Grudin's work [4]); Learn about the taxonomy for CIS systems and situations (e.g., Rodden [8], Golovchinsky [2]) and see how existing systems fit into this structure.
- Hour-2: Discuss the back-end of CIS systems; Look at the works by Pickens et al. [7], Shah et al. [11], and Soulier et al. [12] that focus on algorithmic mediation; Discuss the front-end of collaborative IR systems; Look at the works by Twidale et al. [13] and Morris [6] that focus on interface-based mediation.
- Hour-3: Learn about existing approaches to data mining and evaluation for CIS systems (e.g., Baeza-Yates & Pino [1]); See how system-based and user-based measurements could be used for a more holistic analysis (e.g., Shah & Gonzalez-Ibanez [10]); Understand how prediction and recommendation frameworks could be incorporated in collaborative search (e.g., Gonzalez-Ibanez et al. [3]).

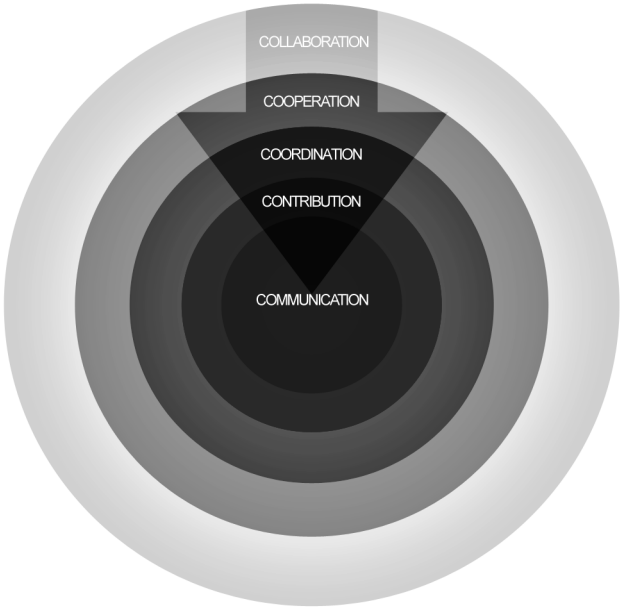

Figure 2: The *C5* model of collaboration.

Following is a more detailed outline of the tutorial presentation.

1. Collaboration - what and why
   (a) Requirement or setup
   (b) Division of labor
   (c) Diversity of skills
2. Defining collaboration - definitions from literature and our own understanding of various terms and terminology
   (a) The *C5* model (see Figure 2)
   (b) Interaction, intention, trust, awareness, etc. in defining collaboration
   (c) CIS in the context of collaboration, information seeking, and IR
3. Background - important people, works, and recent events relating to CIS/CIR
   (a) Various research groups around the world working on CIS/CIR
   (b) Conference events, social gatherings, workshops, etc.
4. Introducing collaboration in IR
   (a) Difference between CIS and CIR
   (b) CIS in the context of collaborative information behavior (CIB), co-browsing, co-search, and collaborative filtering
5. Studying group activities in time and space
   (a) Placement of various collaborative activities on time-space plane
   (b) Role of such placement in studying and/or evaluating CIS systems
6. Extended framework for collaborative activities with 12 different dimensions
   (a) Dimensions include intention, activeness, synchronous-asynchronous, co-located or remote, system-user, awareness, interactiveness, symmetry of roles, strength of the tie, symmetry of benefits, and nature of information activity
7. System-mediated CIS
   (a) Case study: Cerchiamo
   (b) Case study: Querium
8. User-mediated CIS
   (a) Case study: Ariadne
   (b) Case study: SearchTogether
   (c) Case study: Coagmento
9. Discuss generic features of a CIS/CIR system
   (a) Essential features
   (b) Desired features
10. Evaluation using system-focused and user-focused measures (see Figure 3)
    (a) Traditional measures: precision, recall, coverage, novelty, diversity
    (b) Productivity measures: effectiveness, efficiency
    (c) Usability measures: ease of learning, ease of use, satisfaction, cognitive load
    (d) Other: engagement, awareness

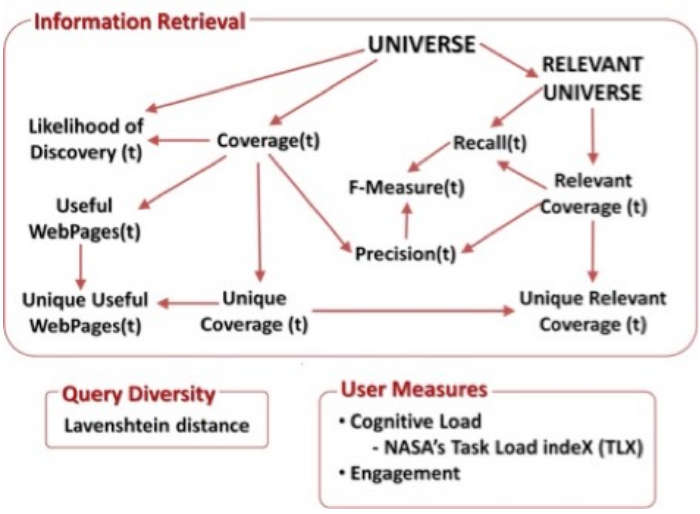

Figure 3: An example of an evaluation framework for collaborative search.

11. Creating synergic solutions for CIS
    (a) How does 1+1 become greater than 2 in collaboration
    (b) Issue of control in the context of system-mediated and user-mediated collaborations
    (c) Achieving selective collaboration to provide a balance of user-control and system-intelligence for synergy
12. Applications of CIS
    (a) Exploratory search (e.g., legal informatics, patent searching, healthcare)
    (b) Education
    (c) Product search (e.g., using selective collaboration method on Amazon and eBay)
    (d) Generalizing recommender systems for recommending people and processes
13. Future directions for research and development

5. COURSE MATERIALS

Course materials will be made available on a Web site hosted under http://collab.infoseeking.org/ two weeks prior to the tutorial. They will include links to relevant software; links to publications that will be discussed; and mechanisms for communication (Twitter hashtags, blog space, etc.) among the tutorial participants, before, during and after the tutorial.

Since this course is focused on introducing several relevant theories, models, approaches, and findings in the field of CIS/CIR, the participants will also be given an annotated bibliography that represents seminal works in this area.

6. REFERENCES

[1] R. Baeza-Yates and J. A. Pino. A First Step to Formally Evaluate Collaborative Work. In S. C. Payne and W. Prinz, editors, *Proceedings of ACM GROUP Conference*, pages 56–60, Phoenix, AZ, 1997.

[2] G. Golovchinsky, J. Pickens, and M. Back. A taxonomy of collaboration in online information seeking. In *Proceedings of JCDL 2008 Workshop on Collaborative Exploratory Search*, Pittsburgh, PA, June 2008.

[3] R. González-Ibáñez, C. Shah, and R. W. White. Capturing collabportunities: A method to evaluate collaboration opportunities in information search using pseudo-collaboration. *Journal of Association of Information Science and Technology (JASIST)*, 2014.

[4] J. Grudin. Groupware and social dynamics: eight challenges for developers. *Communications of the ACM*, 37(1):92–105, Jan. 1994.

[5] P. Hansen, C. Shah, and C.-P. Klas, editors. *Collaborative Information Seeking - Best Practices, New Domains, New Thoughts*. Springer, 2015.

[6] M. R. Morris and E. Horvitz. SearchTogether: An Interface for Collaborative Web Search. In *ACM Symposium on User Interface Software and Technology (UIST)*, pages 3–12, Newport, RI, Oct. 2007. New York, N.Y.: ACM Press.

[7] J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. Algorithmic Mediation for Collaborative Exploratory Search. In *Proceedings of the Annual ACM Conference on Research and Development in Information Retrieval (SIGIR)*, pages 315–322, Singapore, July 2008.

[8] T. Rodden. A Survey of CSCW Systems. *Interacting with Computers*, 3(3):319–353, 1991.

[9] C. Shah. *Collaborative Information Seeking: The Art and Science of Making The Whole Greater Than The Sum of All (The Information Retrieval Series)*. Springer, Berlin, Germany, 2012.

[10] C. Shah and R. Gonzalez-Ibanez. Evaluating the Synergic Effect of Collaboration in Information Seeking. In *Proceedings of the Annual ACM Conference on Research and Development in Information Retrieval (SIGIR)*, pages 913–922, Beijing, China, 2011.

[11] C. Shah, J. Pickens, and G. Golovchinsky. Role-based results redistribution for collaborative information retrieval. *Information Processing & Management*, 46(6):773–781, Nov. 2010.

[12] L. Soulier, C. Shah, and L. Tamine. User-driven system-mediated collaborative information retrieval. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval - SIGIR '14*, pages 485–494, New York, New York, USA, jul 2014. ACM Press.

[13] M. B. Twidale, D. M. Nichols, and C. D. Paice. Supporting collaborative learning during information searching. In *Proceedings of Computer Supported Collaborative Learning (CSCL)*, pages 367–374, Bloomington, Indiana, 1995.