SECC: A Novel Search Engine Interface with Live Chat Channel

SECC: A Novel Search Engine Interface with Live Chat Channel

Cheng Zhang [1], Peng Zhang [*,1], Jingfei Li [1], Dawei Song [*,1,2]
[1]Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, China
[2]School of Computing and Communications, The Open University, United Kingdom
zccode@gmail.com
{pzhang, jingfeili, dwsong}@tju.edu.cn

ABSTRACT

Traditional information retrieval systems rank documents according to their relevance to users' input queries. State of the art commercial search engines (SEs) train ranking models and suggest query refinements by exploiting collective intelligence implicitly using global users' query logs. However, they do not provide an explicit channel for users to communicate with each other in the search process. By asking or discussing with other users on the fly, a user could find relevant information more conveniently and gain a better search experience. In this paper, we present a demo of novel Search Engine with a live Chat Channel (SECC). SECC can group users automatically based on their input queries and allow them to communicate with each other in real time through a chat interface.

CCS Concepts

•Information systems → Search interfaces; •Software and its engineering → *Software design engineering;*

Keywords

Information Retrieval; Search Engine; Collective Intelligence; Chat Channel

1. INTRODUCTION

Users' collective intelligence, such as the collective users' query and interaction history recorded in search logs, has been widely used to support various important features, e.g., query recommendation [1], query completion [3] and webpage ranking [9], etc.

As shown in Figure 1, traditional search engines (TSEs) typically adopt a similar user interface, including a search control, display of search results and related searches. As an example, when a user inputs a query "information retrieval", popular TSEs such as Google, Bing and Baidu, all provide a list of related searches (as query recommendations), illustrated at the bottom of Figure 1.

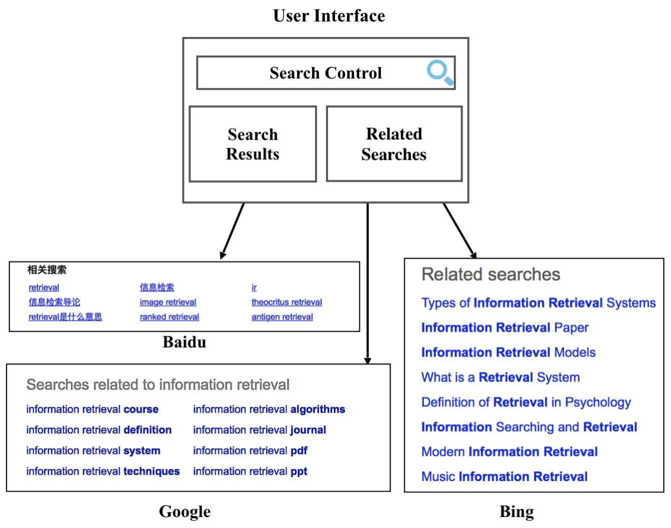

User Interface
Search Control
Search Results
Related Searches
相关搜索
retrieval
信息检索
ir
信息检索导论
image retrieval
theocritus retrieval
retrieval是什么意思
ranked retrieval
antigen retrieval
Baidu
Searches related to information retrieval
information retrieval course
information retrieval algorithms
information retrieval definition
information retrieval journal
information retrieval system
information retrieval pdf
information retrieval techniques
information retrieval ppt
Google
Related searches
Types of Information Retrieval Systems
Information Retrieval Paper
Information Retrieval Models
What is a Retrieval System
Definition of Retrieval in Psychology
Information Searching and Retrieval
Modern Information Retrieval
Music Information Retrieval
Bing

Figure 1: Common functionalities of popular search engines.

The components of query recommendation, query completion and webpage ranking are basic and important functionalities provided by the current search engines. These approaches are mostly trained with global query logs, which can be viewed as an implicit form of collective intelligence. Many search applications have utilized query logs to integrate the wisdom of other users, but the implicit methods could lead to estimation bias in inferring the user's intent [5].

Current TSEs provide little support to the situations where users want to readily communicate with other people to ask questions, especially when dealing with difficult search tasks. Recent attempts on collaborative search systems such as CoFox [4] , CoZapce [6] and SearchTogether [7], tackled this problem to some extent, but most of them focus on the collaboration between small groups of users, e.g., between specific people, colleagues and friends. It can be viewed as a kind of narrow collective intelligence.

In order to exploit global users' wisdom explicitly, we propose to add a live chat component in search engines to allow a user to request other users' help conveniently. In this ar-

[*]Corresponding Authors: Peng Zhang and Dawei Song

SIGIR '16, July 17-21, 2016, Pisa, Italy
© 2016 ACM. ISBN 978-1-4503-4069-4/16/07…$15.00
DOI: http://dx.doi.org/10.1145/2911451.2911454

ticle, we present a novel demo Search Engine with dynamically established live Chat Channels (SECC). SECC provides a socialized search function by implementing a user-friendly online chat interface for users who share similar search queries. Compared with TSEs, SECC allows similar users to communicate with each other to discuss about commonly concerned topics or problems. This new function will potentially benefit users' search experience.

The rest of the paper is organized as follows: In Section 2, we illustrate our motivation. The design and implementation of the demo is introduced in Section 3. Then, we demonstrate the SECC system in Section 4. Finally, Section 5 concludes the paper and discusses about future work.

2. MOTIVATION

In this section, we introduce our motivation from two perspectives: topology of search engine architecture and usage of collective intelligence. We first summarize the overall differences of SECC compared with TSE in Table 1.

	Topology	Collective Intelligence
TSE	single-center	implicit
SECC	multi-center	explicit

Table 1: Comparation between SECC and TSE in terms of topology and the usage of collective intelligence

2.1 Moving Towards a Multi-center Topology

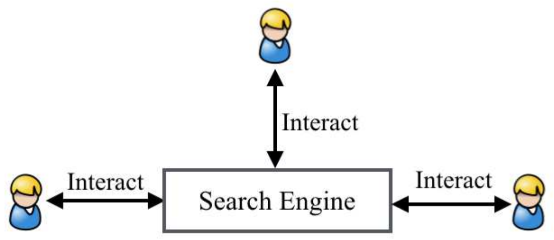


Figure 2: Single-Center interaction topology

Typical user interaction with a TSE can be summarized as follows: after a user submits a query, the SE returns a ranked list of documents to the user. Then, the user clicks on some results without any explicit communication with other people. This is a consequence of the TSE's single-center topology, as shown in Figure 2. In our opinion, TSE has a kind of star-like topology with a single center. All users communicate with the search engine. The feature of star-like topology is that it is easy to add new nodes (new users) into the network, but the failure of the central node will cause a halt of the entire network. If the center fails to provide satisfactory IR services (e.g., query recommendation), the users will be left no other choices to better satisfy their information needs, resulting in a poor user experience.

To address this issue, we propose a novel idea that moves from the single centered structure to a topology with multiple centers. To this end, we add a *Social Engine* to the original topology. As shown in Figure 3, the multi-center topology allows users to interact with the system through multiple channels, called *chat channels* in this paper, through the social engine. As a result, related users can be directly linked and they can instantly communicate with each other during a search session. This would improve the chance that users find desired information when the traditional search engine fails to do so.

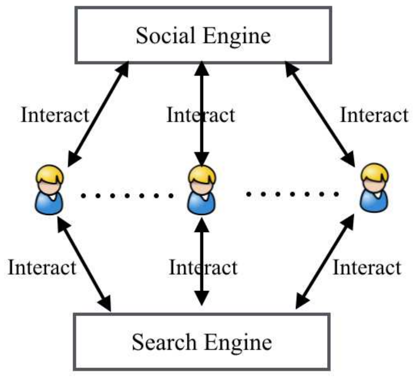


Figure 3: Multi-center interaction topology

2.2 Leveraging Explicit Collective Intelligence

In traditional search engines, collective intelligence has already been considered by mining historical and collective user behaviors, e.g., clicking, reading, query changing, etc. For instance, Boldi et al. [1] construct query flow graphs from global query logs to generate query recommendations. Duan et al. [3] propose to correct spelling errors for query completion based on the distribution of queries in the query log. A webpage re-ranking model is developed with the dynamically grouped users' query log data [9]. These approaches can be viewed as examples of applying implicit collective intelligence in the search process. However, as argued by Clark [2], it is still a challenge for machines to deeply and precisely understand the whole world to answer questions that go beyond the information exactly stated in text. For example, the implicit exploitation of collective intelligence may fail to solve some special problems, such as a mathematical problem described in Table 2 (an exercise that appeared in [8]). Therefore, there is a limitation for the use of implicit collective intelligence in search engines cover all possible search problems.

Consider the functions $f_n(x) = \sin ax$ $(n = 1, 2, 3, ..., -\pi \leq x \leq \pi)$ as points of $\mathscr{L}^2$. Prove that the set of these points is closed and bounded, but not compact.

Table 2: A mathematical problem

To better exploit collective intelligence in SEs, we propose a practical solution by directly adding an explicit collective intelligence component (namely a social engine) into the SECC. The added component provides a live chat channel to a group of users who have similar search tasks. As shown in Figure 3, the dotted lines indicate the links among users. The user grouping should be dynamic with respect to the users' information need. Users in the same group will be provided a channel to communicate online with each other. When users face difficult problems (e.g., the mathematical problem presented in Table 2) which can not be solved by TSE, they can explicitly discuss the problems and share experiences through the chat channel. This would lead to a better chance for users to finally complete the search task through explicit collective intelligence, thus leading to a better user experience.

3. DESIGN AND IMPLEMENTATION

In this section, we design the framework of SECC and introduce its implementation details.

3.1 Framework

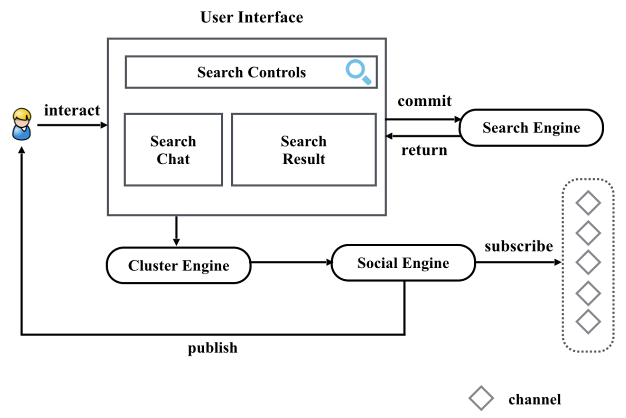


Figure 4: Framework

As shown in Figure 4, our system contains three main engines, i.e., a search engine, a cluster engine and a social engine. The search engine provides traditional functions (e.g., query completion, query suggestion and webpage ranking etc.) based on an indexed corpus. The cluster engine groups similar users dynamically based on overlapping of input queries. The social engine provides a chat channel for grouped users, which allows them to exploit collective intelligence. To facilitate these new functions, we designed a new layout of the user interface (Figure 6 (b)). On the left hand side is the chat panel, while the right hand side displays the search engine results page (SERP) returned by the search engine. The chat function is jointly supported by the cluster engine and the social engine.

When a user inputs a query, the search engine will represent the query with a language model, and then return webpages ranked according to estimated probabilities of relevance between the query language model and the language models of webpages. Meanwhile, the cluster engine will automatically run in background to dynamically group similar users according to users' query terms. Each group will correspond to a chat channel as shown in Figure 4. If a user is not satisfied with the search results, he/she can post questions on the chat channel, and then the social engine will broadcast the questions to all online users who are in the same channel.

It is worth noting that if a user changes his/her input query, the user's group (and chat channel) will change accordingly. The system detects the changes automatically and assists the user to establish another channel in real time.

3.2 Implementation

As shown in Figure 5, the SECC system consists of five modules, including user interface, web server, search engine, cluster engine and social engine.

SECC is coded with Python 2.7.x[1]. We develop the user interface with the bootstrap[2]. The web server is supported by Tornado[3] library which has the property of asynchronous and non-blocking network I/O. A fast full-text indexing and searching library implemented in Python, namely Whoosh[4], is utilized to develop the search engine module. The chat data are stored in the in-memory key-value database Redis[5] and accessible via the Python client for Redis[6]. Additionally, Ajax is used to request and post information for the chat engine. All modules and related technical information are illustrated in Figure 5.

When a user starts a search task, the search engine receives the input queries and return search results by HTTP request. The social engine, instead, uses long-polling to maintain the connection for users all the way through the search session and provides the users with necessary communication services.

In this version of demo, we focus on verifying the feasibility of the proposed framework. Therefore, we do not crawl webpages from the Internet. Instead, we use the TREC AP8889 as test corpus. In the future, we will enlarge the corpus by crawling webpages from the Internet and selected domains to make a more practical system.

4. DEMONSTRATION

We illustrate the demo in Figure 6. Figure 6 (a) shows the start page of our demo, which is in a typical search box style as used in the prominent TSEs. After we input a query into the search box, the system will detect the current user's search interest and cluster a set of users who shares similar search topics. Accordingly, the interface will change to another layout as shown in Figure 6 (b). The chat module is placed in the left side of the main page, while the search results are displayed in the right side (Figure 6 (b) and (c)). The user can communicate with others through live chat, which increases the interactivity among users.

An illustrative example is shown in Figure 6(b) and (c). Suppose two users, with nicknames 'zhang' and 'tju' respectively, share the same query words "Prime (Lending) Rate Moves". SECC will cluster them into the same chat channel, in which they can communicate conveniently. Similarly, if they both want to solve the mathematical problem mentioned in the previous section, they can discuss how to solve it through the chat channel and find the correct answer together. Note that, a user's chat channel will change dynamically in accordance with the change of the user's query words.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we present a novel search engine with a live chat channel, which allows users sharing similar search

[1]https://www.python.org
[2]http://getbootstrap.com
[3]http://www.tornadoweb.org/en/stable
[4]http://bitbucket.org/mchaput/whoosh/wiki/Home
[5]http://redis.io
[6]http://pypi.python.org/pypi/redis

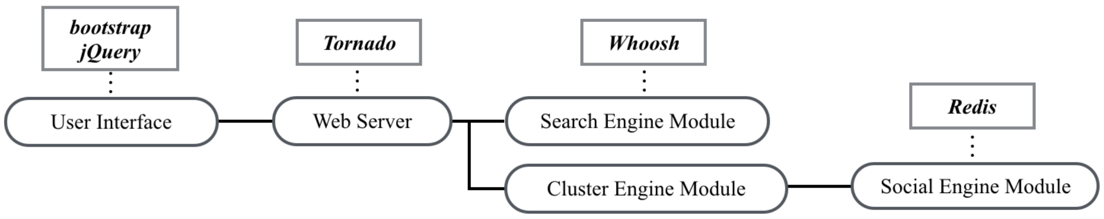
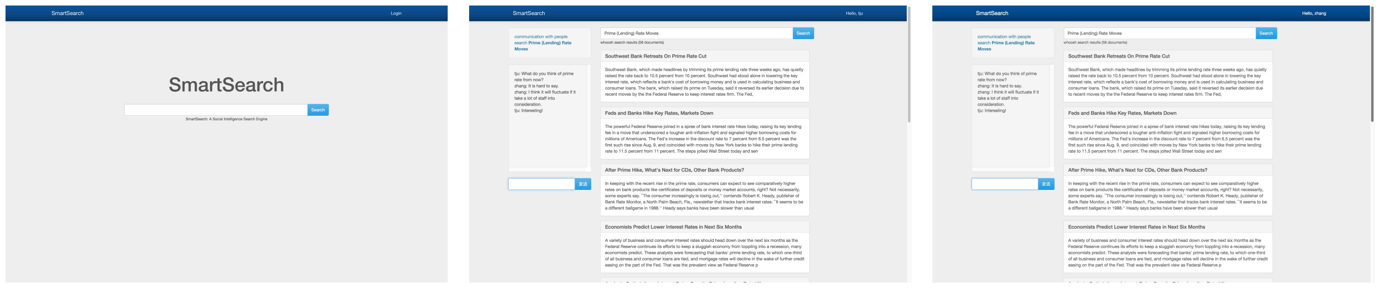


Figure 5: Modules and related technical dependencies



SmartSearch Hello, tju

communication with people search **Prime (Lending) Rate Moves**

tju: What do you think of prime rate from now?
zhang: It is hard to say.
zhang: I think it will fluctuate if it take a lot of staff into consideration.
tju: Interesting!

发送

Prime (Lending) Rate Moves Search
whoosh search results (58 documents)

Southwest Bank Retreats On Prime Rate Cut

Southwest Bank, which made headlines by trimming its prime lending rate three weeks ago, has quietly raised the rate back to 10.5 percent from 10 percent. Southwest had stood alone in lowering the key interest rate, which reflects a bank's cost of borrowing money and is used in calculating business and consumer loans. The bank, which raised its prime on Tuesday, said it reversed its earlier decision due to recent moves by the the Federal Reserve to keep interest rates firm. The Fed,

Feds and Banks Hike Key Rates, Markets Down

The powerful Federal Reserve joined in a spree of bank interest rate hikes today, raising its key lending fee in a move that underscored a tougher anti-inflation fight and signaled higher borrowing costs for millions of Americans. The Fed's increase in the discount rate to 7 percent from 6.5 percent was the first such rise since Aug. 9, and coincided with moves by New York banks to hike their prime lending rate to 11.5 percent from 11 percent. The steps jolted Wall Street today and sen

After Prime Hike, What's Next for CDs, Other Bank Products?

In keeping with the recent rise in the prime rate, consumers can expect to see comparatively higher rates on bank products like certificates of deposits or money market accounts, right? Not necessarily, some experts say. ``The consumer increasingly is losing out,'' contends Robert K. Heady, publisher of Bank Rate Monitor, a North Palm Beach, Fla., newsletter that tracks bank interest rates. ``It seems to be a different ballgame in 1988.'' Heady says banks have been slower than usual

Economists Predict Lower Interest Rates in Next Six Months

A variety of business and consumer interest rates should head down over the next six months as the Federal Reserve continues its efforts to keep a sluggish economy from toppling into a recession, many economists predict. These analysts were forecasting that banks' prime lending rate, to which one-third of all business and consumer loans are tied, and mortgage rates will decline in the wake of further credit easing on the part of the Fed. That was the prevalent view as Federal Reserve p

SmartSearch Hello, zhang

communication with people search **Prime (Lending) Rate Moves**

tju: What do you think of prime rate from now?
zhang: It is hard to say.
zhang: I think it will fluctuate if it take a lot of staff into consideration.
tju: Interesting!

发送

Prime (Lending) Rate Moves Search
whoosh search results (58 documents)

Southwest Bank Retreats On Prime Rate Cut

Southwest Bank, which made headlines by trimming its prime lending rate three weeks ago, has quietly raised the rate back to 10.5 percent from 10 percent. Southwest had stood alone in lowering the key interest rate, which reflects a bank's cost of borrowing money and is used in calculating business and consumer loans. The bank, which raised its prime on Tuesday, said it reversed its earlier decision due to recent moves by the the Federal Reserve to keep interest rates firm. The Fed,

Feds and Banks Hike Key Rates, Markets Down

The powerful Federal Reserve joined in a spree of bank interest rate hikes today, raising its key lending fee in a move that underscored a tougher anti-inflation fight and signaled higher borrowing costs for millions of Americans. The Fed's increase in the discount rate to 7 percent from 6.5 percent was the first such rise since Aug. 9, and coincided with moves by New York banks to hike their prime lending rate to 11.5 percent from 11 percent. The steps jolted Wall Street today and sen

After Prime Hike, What's Next for CDs, Other Bank Products?

In keeping with the recent rise in the prime rate, consumers can expect to see comparatively higher rates on bank products like certificates of deposits or money market accounts, right? Not necessarily, some experts say. ``The consumer increasingly is losing out,'' contends Robert K. Heady, publisher of Bank Rate Monitor, a North Palm Beach, Fla., newsletter that tracks bank interest rates. ``It seems to be a different ballgame in 1988.'' Heady says banks have been slower than usual

Economists Predict Lower Interest Rates in Next Six Months

A variety of business and consumer interest rates should head down over the next six months as the Federal Reserve continues its efforts to keep a sluggish economy from toppling into a recession, many economists predict. These analysts were forecasting that banks' prime lending rate, to which one-third of all business and consumer loans are tied, and mortgage rates will decline in the wake of further credit easing on the part of the Fed. That was the prevalent view as Federal Reserve p

(a) (b) (c)

Figure 6: Demonstration of SECC

topics to interact with each other during search sessions and exploit collective intelligence explicitly. The system provides a live communication service besides traditional IR components.

The current version of the demo is for testing our initial ideas and has a large room for further improvements. Currently, the user cluster engine groups similar users based on overlapping of input queries. This query-based approach is insufficient for understanding the user's intent during a search task (or a session) and would cause too frequent updates of users' chat channels. Thus we will consider grouping users at the session/task level. Secondly, the size of indexed corpus is currently small. We should enlarge the corpus significantly by crawling webpages from the Internet (for selected domains) regularly, in order to make the search engine more practical. This will also attract more users to use our system, so as to improve our system by exploiting the collective intelligence explicitly and implicitly. Furthermore, we will investigate effective mechanisms, such as gamification, to promote the users' incentive to use the chat channels. Finally, we will carry out large scale user task-based evaluation and keep refining our interfaces and algorithms.

6. ACKNOWLEDGEMENTS

The work presented in this paper is sponsored in part by the Chinese National Program on Key Basic Research Project (973 Program, grant No. 2013CB329304, 2014CB744604), the Chinese 863 Program (grant No. 2015AA015403), the Natural Science Foundation of China (grant No. 61402324, 61272265), and the Tianjin Research Program of Application Foundation and Advanced Technology (grant No. 15JC-QNJC41700)..

7. REFERENCES

[1] P. Boldi, F. Bonchi, C. Castillo, D. Donato, A. Gionis, and S. Vigna. The query-flow graph: model and applications. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 609–618. ACM, 2008.

[2] P. Clark. Elementary school science and math tests as a driver for ai: Take the aristo challenge. *to appear*, 2015.

[3] H. Duan and B.-J. P. Hsu. Online spelling correction for query completion. In *Proceedings of the 20th international conference on World wide web*, pages 117–126. ACM, 2011.

[4] R. P. Jesus, T. Leelanupab, and J. M. Jose. Cofox: A synchronous collaborative browser. In *Information Retrieval Technology*, pages 262–274. Springer, 2012.

[5] T. Joachims and F. Radlinski. Search engines that learn from implicit feedback. *Computer*, (8):34–40, 2007.

[6] H. Kruajirayu, A. Tangsomboon, and T. Leelanupab. Cozpace: a proposal for collaborative web search for sharing search records and interactions. In *Student Project Conference (ICT-ISPC), 2014 Third ICT International*, pages 165–168. IEEE, 2014.

[7] M. R. Morris and E. Horvitz. Searchtogether: an interface for collaborative web search. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 3–12. ACM, 2007.

[8] W. Rudin. *Principles of mathematical analysis*, volume 3. McGraw-Hill New York, 1964.

[9] T. T. Vu, D. Song, A. Willis, S. N. Tran, and J. Li. Improving search personalisation with dynamic group formation. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 951–954. ACM, 2014.