

ソーシャルクラウド型知識情報獲得支援システムの研究開発について

秋山 豊和[†] 河合 由起子[†] 松井 優也[†]
 Toyokazu Akiyama[†], Yukiko Kawai[†], Yuya Matsui[†]
[†]京都産業大学
[†]Kyoto Sangyo University

我々はこれまでにハイパーリンクとソーシャルリンクを用いることで、情報検索およびソーシャルコミュニケーションの双方の利点を生かした検索システムを提案している。本稿ではこれまでに提案したシステムについて紹介し、提案システムの実用化に向けた問題点と今後の開発方針について触れる。

1. はじめに

知識獲得のため Web 検索を行う場合、2つの方法が考えられる。一つは、検索サービスの利用であり、もう一つは、掲示板や SNS の利用である。検索サービスは速度と網羅性の点で有効であり、SNS を利用した情報収集は検索サービスよりも時間や手間は要するが、コミュニケーションにより質の高い情報が得られるという利点がある。しかし、両方の利点を同時に活用したサービスはいまだない。近年では、ページの情報から専門家を抽出し、検索キーワードに対する専門家を特定し、その専門家の連絡先をユーザへ提供したり、複数人で協調して検索するといった研究開発も行われている [1][2]。

我々はこれまで、検索と SNS 双方の利点を同時に生かせるサービスとして、Web ページをライブで閲覧しているユーザのネットワークを構築することで、1) 閲覧者の量と質に基づいたランキングならびに2) ページを通して閲覧者と即時的なコミュニケーションを実現することを旨としたシステムを提案している [3][4]。検索手法は Pagerank アルゴリズムに基づいており、各ページ間のリンクに対して閲覧者のアクセス数とアクセス時間を考慮した重みを付与する。既存の検索エンジンから得られた結果の各ページは、この重みが付与されたリンク構造によって再ランキングされ、ユーザは検索結果の中から閲覧者が最も多いページを獲得できる。これにより、「瞬時に欲しい情報」と「瞬時に検索時に同じ興味を持つ、あるいは知識を豊富に持つユーザからの支援」を同時に獲得できる。

本稿では、提案システムの概要および利点について示し、現状での問題点と今後の開発方針について示す。

2. システム概要

図1にシステムの概要を示す。本稿で提案するシステムはユーザから検索キーワードを取得すると、リアルタイム閲覧者の質と量に基づいて順位付けした検索結果を提示する。また、検索結果として表示される各リンクには今現在の閲覧者数が表示される。ユーザはリンクをクリックするとページの閲覧と共に、そのページを閲覧している他ユーザとその場でチャットによるコミュニケーションを取ることができる。このコミュニケーションログは各ページごとに公開される。さらに、ページ内の全

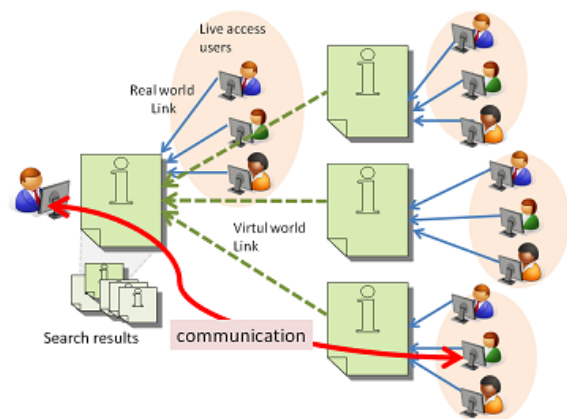


図 1: システムの概要

てのリンクにも閲覧者数が表示される。これにより、検索結果ページを通して、より深いリンク先のページにいるユーザともコミュニケーションでき、検索キーワードと関連の深い/浅い、ユーザからも支援が受けられる。

2.1 閲覧人数に着目した検索

提案システムでは Web ページの閲覧人数に着目した検索を提供する。既存の検索サービスにおけるページの重要度というものは、被引用数であるリンクの数に着目したアルゴリズムによって決定付けられる。その精度は年々向上しているが、検索結果には閲覧者の評価が十分に反映されているとは言えない。そこで提案システムでは、従来の PageRank アルゴリズムのリンク構造による検索手法を用いて、ページ閲覧者の質と量に基づいてリンクに重み付けすることで、今この瞬間における実際のユーザ評価に基づいた検索を提供する。

さらに提案システムでは、全ての Web ページの中で最も閲覧者数が多いページ群の提示や、閲覧中の Web ページのリンク先にいるユーザの数を提示することが可能となるため、他ユーザが参考にしているページをより多く発見できる。

2.2 コミュニケーション機能

提案システムでは Web ページをチャット・ルームに見立て、同一ページを閲覧している複数のユーザ同士でキャラクタを用いた簡単なコミュニケーションを行うことができる (図2)。コミュニケーション機能は提案システムのログインボタンを押すことにより開始される。



図 2: コミュニケーション画面

コミュニケーション機能が開始されるとライブ閲覧者のキャラクタと右に会話用ウィンドウが表示される。会話用ウィンドウ下部の入力フォームを用いることで、同一ページの閲覧者に対して発話できる。Web ブラウジングの合間にページ内の不明な点の問い合わせや、気楽なコミュニケーションができる。発言内容はサーバの発話ログ内に格納され、時間を跨いで参照可能である。

また、話題や疑問の対象にしたい文章をハイライトし、同一ページの他閲覧者と共有できる。これにより、ページ中の疑問点や話題点を他閲覧者に容易に気付かせることができ、円滑にコミュニケーションができる。尚、受け取るユーザ側は拒否することもできる。

3. システムの有効性に関する考察

Web 上のサイトやページの信頼度や評価を計る手法として、閲覧者数やブックマーク数などのユーザ評価を用いたアプローチは数多く施行されており、各種ブログサービスの人気度ランキングや近年盛況してきているソーシャルブックマークによるページ評価もその一つである。一方、本システムでは閲覧者数でのランキングによる検索精度の向上や信頼できるページの発見もさることながら、リアルタイムな閲覧者数の表示と、あらゆる場所をコミュニケーション可能な場として提供することで、今すぐコミュニケーション可能なユーザの発見を促している。これによってより閲覧者同士のコミュニケーションによる情報収集と交流を実現を重視した形となっている。また、同一ページを閲覧したユーザ同士によるコミュニケーション手法は、古くからは Web サイトごとに設けられた掲示板やコメント欄があるが、本システムではそのようなコミュニケーションのための仕組みが用意されていないページにおいてもコミュニケーションを可能としている。例えば Google で「Java」と検索して返される検索結果のページの中でさえもコミュニケーションが可能である。ページごとのユーザ同士の交流によりどの程度有益な情報が手に入るかどうかは未知数で

あるが、Web ブラウジング中に、いつでもどこでもユーザが発見可能であり、かつコミュニケーションを可能とすることによって、Web ページがより活発な情報収集と交流の場として機能することを期待している。

4. 提案システムの実用化に向けた拡張

提案システムでは、Firefox の XUL (XML User Interface Language) を用いてブラウザの plugin として閲覧人数に着目した検索機能およびコミュニケーション機能を実装している。コミュニケーション機能では、Tomcat Web アプリケーションサーバで提供されている Comet により非同期通信を実現している。また、閲覧者数に着目した検索では、ユーザが指定したクエリを Yahoo! Japan の Web API で検索し、取得した結果を DB 上の閲覧者数データによってリランキングしている。この際、サーバ上で PageRank 計算のための固有値計算を行っている。提案システムは现阶段はシステム評価のためのプロトタイプであり、リランキングの計算性能が低く、システムの実用化に向けて、高速化する必要がある。固有値計算においては、Lanzcos, Arnoldi 等の手法を用いて並列化し、高速化を図る。計算環境の構築には JGN2plus の仮想化サービスを利用し、さらに、安価にプロセッサ数を確保するため、GPGPU による並列計算を併用する予定である。一方で、リアルタイムに閲覧しているユーザをコミュニケーション対象として捉えるためには、非常に短時間で閲覧ユーザの影響を検索結果に反映する必要がある。単純な高速化だけでなく、多少精度を犠牲にしてもリアルタイムに結果を返す手法も検討する必要がある。また、コミュニケーション機能の性能評価も行う予定である。

5. まとめ

本稿ではこれまでに提案しているシステムを実用化に向けて拡張したソーシャルクラウド型知識情報獲得支援システムの開発指針について述べた。今後はシステムの実装を進め、実環境における評価を行っていく予定である。

謝辞

本研究の一部は、SCOPE 若手 ICT 研究者育成型研究開発の助成を受け実施したものである。

参考文献

- [1] E. Y. Chang. Confucius and “Its” Intelligent Dis-ciples. In *Proc. CIKM2009*, 2009.
- [2] J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. Algorithmic Mediation for Collaborative Exploratory Search. In *Proc. SIGIR2008*, pp.315-322, 2008.
- [3] 松井 優也, 河合 由起子: 閲覧者ネットワークによる情報収集支援サービスの提案, *WebDB Forum*, (2009).
- [4] Wired: <http://klab.kyoto-su.ac.jp/~mito/>