

情報整理を支援する対話型検索インタフェースの提案と評価

林 大策[†] 福原 知宏^{††} 佐藤 哲司[†]

[†] 筑波大学図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

^{††} 産業技術総合研究所サービス工学研究センター 〒135-0064 東京都江東区青海 2-3-26

E-mail: [†]hayashi@slis.tsukuba.ac.jp, ^{††}tomohiro.fukuhara@aist.go.jp, ^{†††}satoh@ce.slis.tsukuba.ac.jp

あらまし Web ページの増加は著しく、知りたい情報をその位置づけや周辺情報まで含めて調べることは、いっそう難しくなっている。本論文では、従来のキーワード入力型情報検索モデルを拡張し、検索結果の可視化と対話的な操作をシームレスに連動させた対話型検索インタフェースを提案する。ユーザは、検索結果から抽出されたキーワードを用いて検索結果を 2 次元平面上に整理・配置する操作を繰り返しつつ、新たな検索の観点を与えることができる。提案インタフェースを実装しコンセプトマップを用いた利用者実験を行ったところ、特徴的な検索サイクルや探索した話題範囲の広がりがみられた。

キーワード 情報検索, インタラクティブ, 可視化, インタフェース, コンセプトマップ

Design and Implementation of Interactive Information Retrieve Interface to Support Information Rearranging

Daisaku HAYASHI[†], Tomohiro FUKUHARA^{††}, and Tetsuji SATOH[†]

[†] Graduate School of Library Information and Media Studies, University of Tsukuba
1-2, Kasuga, Tsukuba, Ibaraki, 305-0855 Japan

^{††} Center for Service Research, National Institute of Advanced Industrial Science and Technology
2-3-26, Aomi, Koto, Tokyo, 135-064 Japan

E-mail: [†]hayashi@slis.tsukuba.ac.jp, ^{††}tomohiro.fukuhara@aist.go.jp, ^{†††}satoh@ce.slis.tsukuba.ac.jp

Abstract Due to the rapid increase of web pages, information search, including its around information, becomes more difficult. In this paper, we propose a attractive search interface, which combines search-result presentation and user interaction. Users can easily rearrange the iconized pages in two dimensional surface by using sofisticated keywords which is extracted from search result pages. We evaluate a experiment that compared proposal system with popular search system with using concept map. From the results, we confirmed part of effectiveness about the improvement of the search cycle.

Key words Infomation Retrieval, Interactive, Visualization, Interface, Concept Map

1. はじめに

近年は情報爆発と言われるように、インターネット情報空間は爆発的に文書数が増え続けている。そのため、少ない検索語で検索を行うと膨大な量の情報が見つかり、ユーザはそこから情報の取捨選択を行わなくてはならない。また、ユーザが検索で求める情報は多様化しており、多くの検索エンジンは様々な観点の情報を上位に提示するようになってきている。そのため、ユーザがよく目にする検索結果の上位には、別々の話題が混在している場合が多い。たとえば、あるニュースについて検索をしたとき、そのニュースの速報や解説記事、感想などが上位に提示されたとする。そこで“感想”に注目し「他にも“感想”に

関するページはないか」という要求を持ったとしても、その文書の近くにある文書は解説記事やまとめ記事など、別の観点の文書である可能性が高い。ユーザ“感想”に関する文書を見つけるためには、検索結果のページを更にスクロールしていくか、検索質問を修正するなどの方法が必要になる。前者では、スクロールによって確実に文書を見つけられるかは不明である。後者では、追加する検索語がすぐに想起できないなど、検索質問の修正には困難が伴う。追加した検索語が曖昧なものであれば、検索結果の上位には様々な話題が混在してしまう可能性もある。このように、ある情報や文書を発見しても、その情報の位置づけや関連情報などの周辺的な情報を整理・発見していくことは難しくなっている。

本論文では、観点ごとに文書をまとめる検索結果の提示と、対話的な観点の変更とをシームレスに融合した検索インタフェースを提案する。従来は1次元であった検索結果の提示を、2次元に可視化することで情報の発見を容易にし、可視化された文書の配置を対話的に変更することによって、位置づけや周辺の情報を整理することを目指す。これによりユーザの検索の手間を軽減し、情報発見と整理を支援することが本研究の目的である。

実装した提案システムと従来の検索システムとを比較する、利用者実験を行った。その際、情報の整理にはコンセプトマップを用い、課題で調べた情報のまとめを行った。各システムでのマップの広がり方、検索の利用法などを比較し、その違いを明らかにする。

以下、2章では関連研究を挙げ、3章では提案手法の概略について、4章では実験とその結果について述べ、5章で結果を考察し、6章でまとめと今後の課題について述べる。

2. 関連研究

2.1 検索結果の可視化

多くの文書が混在する検索結果を分かりやすく閲覧するのに可視化は有効な手法であり、多くの先行研究がある。林ら [1] は、マトリックス上に検索結果のキーワードを並べる可視化手法を提案している。Robertsら [2] は、検索結果を主にリンク構造に注目して2次元平面上に並べる可視化手法を提案している。小林ら [3] は、検索結果をクラスタリングし、内容が類似するページを近くに配置するように検索結果を可視化している。神原ら [4] は、Web上での商品購入のために、値段や売れ筋などの軸を設定して商品画像を2次元平面上に可視化している。他にも、パネモデルを利用した検索結果の可視化や3次元空間への可視化など、多くの提案・研究が行われている。しかし、実際に利用者による実験を行い、従来の検索との比較をしているものは少ない。本研究では可視化インタフェースを利用者実験によって評価する。可能な限り可視化した情報をわかりやすく提示するため、2次元平面上にページのアイコンを置く検索結果の可視化手法を用いる。

2.2 可視化情報を用いた対話的な検索

グラフィカルで対話的な操作による、結果の再ランキングを行うインタフェースについての研究も多数行われている。吉田ら [5][6] は、検索結果を可視化して閲覧するのではなく、検索結果に出現するキーワードの関係などを可視化し、再ランキングの操作を行える手法を提案している。キーワードをノードとして配置し、ノードの位置を変更することで検索結果を再ランキングする手法である。中谷ら [7] は、検索結果中から特徴的なキーワードを取り出してレーダーチャートを作成し、各項目の重みを変更することで検索結果を再ランキングするインタフェースを提案している。これらの研究は検索結果となるページの間接関係を直接可視化するのではなく、ページ間の関係を部分的に可視化しているといえる。本論文で提案する検索インタフェースはページの間接関係を直接可視化しながら操作を行い、検索結果の再ランキングを行う。そのため、再ランキングした結果の変化についても分かりやすくなると考えられる。

2.3 検索で得る情報の評価

従来の検索インタフェースと提案インタフェースで利用者による比較を行う場合、どのように結果が異なるかを評価することは一般的に難しい。データセット中の適合文献をどれだけ多く見つけられるかという実験は多くあるが、実際のWebを検索して比較を行う場合、適合文献を全て把握することは不可能である。

江草ら [8][9] は、コンセプトマップを用いて、検索の前後でのユーザの知識の変化を計る手法を提案している。これらの研究では、検索の前にコンセプトマップを利用者が作成し、検索の後で再びコンセプトマップを作成し、違いを確認している。本研究では、検索をしながら逐次コンセプトマップを作成することでその違いを確認する手法を用いる。各検索システムを使う前にコンセプトマップを作成し、検索をしながら続きのマップを作成する。

3. 提案手法

本論文では、検索結果中のページの間接関係を分かりやすく可視化し、対話的に操作が可能なインタフェースを提案する。本章では提案インタフェースの概要や、検索結果の可視化手法などについて述べる。

3.1 提案手法の概要

検索結果の可視化は、2つのキーワードを軸として、2次元平面上にページを配置する。縦軸と横軸にキーワードを設定することで、ページの座標から内容を推測しやすくするためである。

提案手法における可視化の方針と、対話的な検索操作の概念を図1で示す。図1(a)は、横軸にキーワード「A」を、縦軸にキーワード「B」を設定した場合の結果であり、縦軸のキーワード「B」を、軸キーワード候補として提示された「 \square 」と置き換えた場合を示したものが図1(b)である。図の(a)、(b)いずれにおいても、平面上に配置したページ①～⑥は、ある検索質問Qによって得られた6件の検索結果ページのアイコンであり、軸キーワードを入れ替えることで同一のページ集合が再配置されることを表している。いずれの図においても、右側にあるページほど横軸キーワードとの関連度を高く、上側にあるページほど縦軸キーワードとの関連度が高いとする。ユーザがQを与えて図1(a)の結果を得た時、ページ④は、二次元平面の上端にあることからキーワードBと関連が強いと判断できる。また、ページ⑤がAというキーワードと関連が高いことも分かる。A,Bを含む軸キーワード候補 \square 、 \square はQによって得られた検索結果から抽出したものである。ここでユーザが、縦軸キーワードをAから \square に置き換えることで、図1(b)の結果が得られたとする。新たに設定された縦軸キーワードとページ間の関連度が再計算され、各ページの配置が変化する。このとき計算するのは縦軸との関連度のみで、横軸については更新しない。そのため全てのページのx座標は変化しない。新たに得られた結果を見ると、キーワード \square と関連の強いページは①と⑥であることが図1(b)の結果から分かる。このように、ユーザが軸キーワードを取り換えることで、ダイナミックに検索結果が再配置され、ページの分布形状から軸キーワードとの

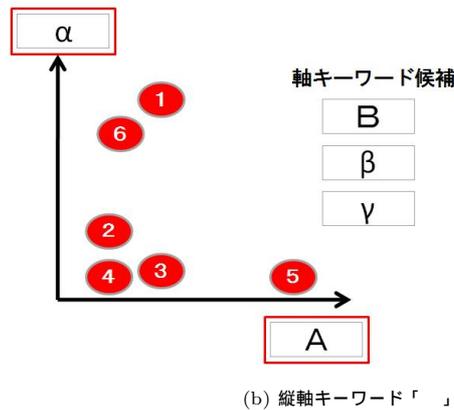
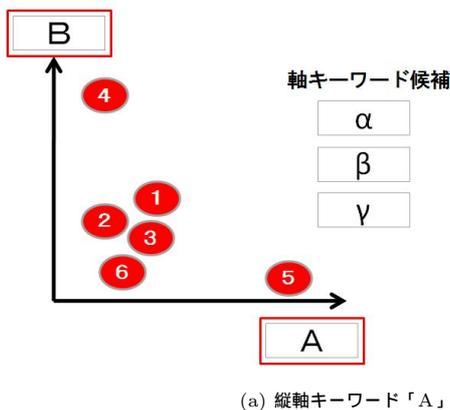


図 1 検索操作の概念

関係の強さや、結果ページ間の関係性などの特徴を知ることができ、目的の情報を見つけやすくなると期待される。

3.2 提案インタフェースの実装

本節では実装したインタフェースについて述べる。図 2 は実装したインタフェースの検索画面である。画面の左上に検索質問の入力フォームがあり、件数を指定してから検索を行う。図の例は、検索質問「TPP」を入力して得られた 60 件の検索結果の配置を示している。

図中の「1」と書かれた領域には、60 件のページがアイコンで表示されている。「2」と書かれた左右の領域には、検索結果ページ集合から抽出した 15 件ずつキーワードが提示されている。「1」の領域の縦軸と横軸には、「2」の領域から選んだキーワードがそれぞれ設定できる。設定された横軸キーワードに関連の近いページは平面上の右側に集まり、縦軸キーワードに関連の近いページは平面上の上側に集まる。図 2 の検索結果では、横軸に「議論」という単語が設定されている。そのため、平面上の右側には「議論」に関連する情報が現れている。TPP に関する記事の中でも、議論について注目したページが見たい場合などに選ぶと効果的である。縦軸には「農業」という単語が設定されている。平面上の上側には「農業」に関連するページが現れる。ユーザはシステムが提示したそれぞれのキーワードから任意に選んで設定することができる。

「3」の領域には 5 つのボックスがあり、これらは平面上で対応する番号のアイコンのタイトルやスニペットを表している。平面上には 3 つの色のアイコンが存在する。青色のアイコンは、平面上でのページの位置を示している。水色のアイコンは、現在平面上で注目されている部分を示している。図 2 では平面上の上側に 5 つの水色のアイコンが見える。この水色になっているアイコンに対応するページの内容が「3」のボックスに表示される。平面上のいずれかの場所をクリックすることで、クリックした箇所の近傍のアイコンに注目することができる。図 2 の縦軸キーワードは「農業」になっており、「3」の領域で確認すると、各ページのタイトル・スニペットに「農業」という単語が出現していることが分かる。これは「農業」と関連のあるページが上側に出てきているということである。また、赤色のア

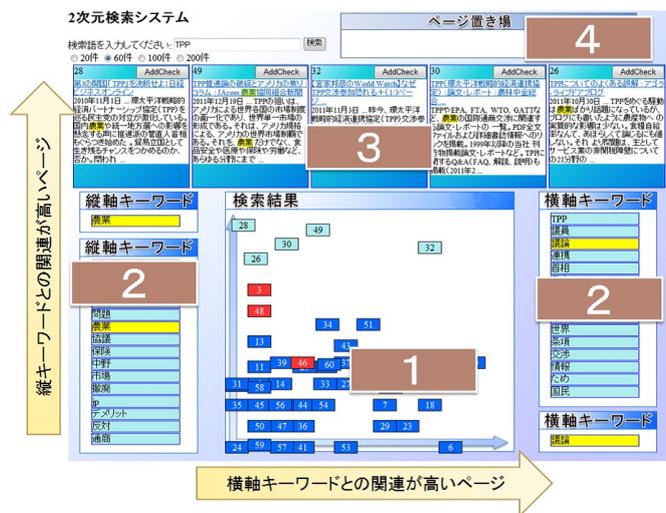


図 2 実装したインタフェース

イコンは最後に選んだ軸キーワードを含むページである。図 2 は縦軸キーワードに「農業」を設定した後の画面であり、「農業」という単語を含むアイコンが平面上で赤くなっている。この赤色のアイコン近辺をクリックすることで、更に「農業」に関連するページを見ることができる。赤いアイコンの数で「農業」という単語を持っているページがいくつ検索結果中にあるかすぐに判断することができるので、更に情報が欲しい場合に無駄な探索をせず、すぐに検索語に追加をすることができる。

「4」には簡易的なお気に入り機能があり、「3」のボックスで AddCheck と書かれたボタンを押すか、直接ドラッグアンドドロップすることでこのエリアにページを置くことができる。画面の更新や新たな検索を行った場合でも、保存されたページは残っている。

3.3 処理の流れ

図 3 は実装したシステムの処理の流れである。検索時の流れと、軸キーワード選択時の流れを表現している。図 3 を基にして、実装したシステムについてを説明する。

3.3.1 検索の流れ

図 3 には赤の番号と灰色の番号の流れが存在している。赤の

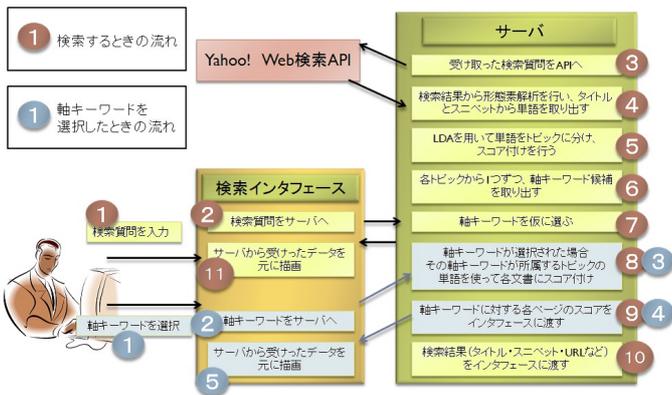


図3 実装システムの処理の流れ

番号は、検索質問の入力をしたときの処理の流れを表している。灰色の番号は、軸キーワードを選択した場合の処理の流れを表している。ここでは赤色の1~11で検索の流れを説明する。

最初にユーザは検索質問をインタフェース上で入力し、その検索質問をサーバに渡す。サーバは、それを検索エンジンに渡し、その結果を受け取る。検索エンジンはYahoo!検索 Web API^(注1)を利用した。返却された検索結果を加工し、インタフェースに返却するという流れになる。以降の節では、3の各フェーズについて説明をしていく。

3.3.2 キーワードの抽出

図3の4に対応するフェーズである。このフェーズでは検索結果の加工や軸キーワード候補を選ぶために単語の抽出を行う。

ユーザが選択する軸キーワードの候補を抽出する手順は以下の通りである。まず、検索結果のページからタイトルとスニペットを対象として形態素解析を行なう。形態素解析ソフトにはMeCab^(注2)を使用して、「一般名詞」「固有名詞」「形容詞」を抽出する。「代名詞」「記号」「数詞」「助数詞」「感動詞」「動詞」などは、ページを特徴づけるキーワードとしては不適切と考え除外した。形容詞を抽出するのは、「軽い」「赤い」「安い」携帯電話など、ページのに記述された対象の特徴を表現するとき形容詞が多用されていることを考慮したからである。上述した方法で生成されたキーワードに得点を付けることで、軸キーワードとして有効な候補を抽出する。

3.3.3 キーワードの得点付け

3の5,6,7に対応するフェーズである。3.3.2で抽出したキーワードを用いてLDA(Latent Dirichlet Allocation)の計算を行う。なお実装にはインターネットで公開されているコード^(注3)を使用した。LDAで軸キーワード候補の数だけ潜在トピックを作成し、各トピックの出現確率が高い上位50語に得点を与える。なお、得点は出現確率をそのまま付与する。各トピックで最も出現確率が高い語を代表とし、全てのトピックで代表語が重複しないように調整を行う。その結果を軸キーワード候補とする。その中から暫定的に縦軸となるキーワードと横軸となる

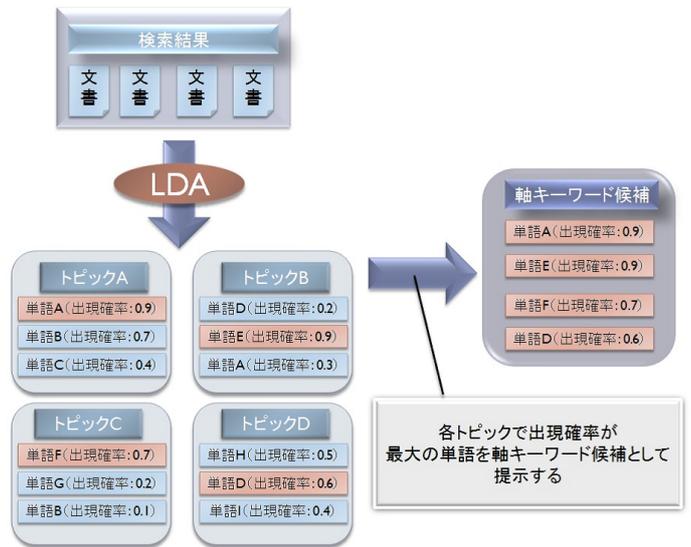


図4 軸キーワード候補提示の流れ

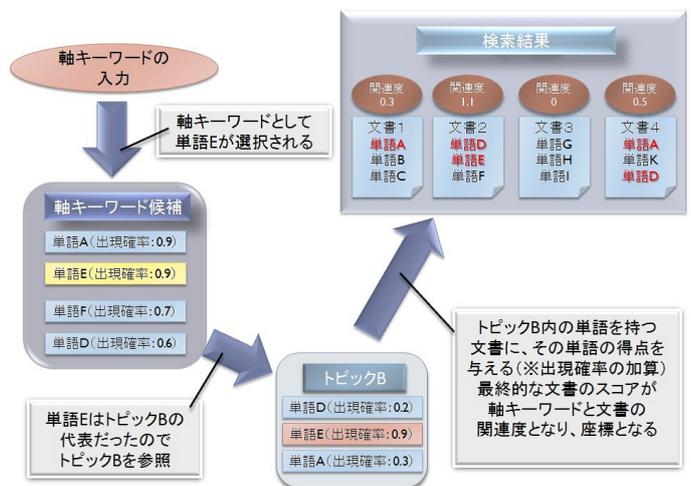


図5 関連度計算の流れ

キーワードを選び、検索質問に対する初期の可視化結果を生成する。図4はこの流れを図で表現している。

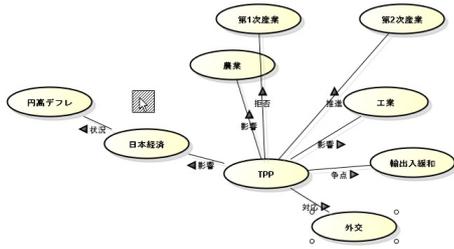
3.3.4 関連度計算による配置方法

図3の8,9,10,11に対応するフェーズである。7のフェーズで決まった、あるいはユーザから入力された軸キーワードを起点に各文書の座標を求める。具体的には、軸キーワードが所属するトピックの単語群を用いる。それらの単語集合を一時的に得点付与語とする。得点付与語を1つずつ、各文書に出現するかどうかをチェックしていく。もし、ある文書が得点付与語のいずれかを持っている場合、その得点付与語のスコア、つまりトピックでの出現確率を付与する。全ての得点付与語を用いて、全ての文書にスコアを付与していき、最終的に全ての文書にスコアを付ける。それらのスコアを0~1で正規化して、その値を軸キーワードに対する座標とする。このようにして1軸での座標を求め、2つの軸キーワードで座標が決まったときに平面上での座標が決定する。その後、各文書のデータと軸キーワード候補集合をインタフェースに返却し、描画を行う。図5はこの流れを図で表現している。

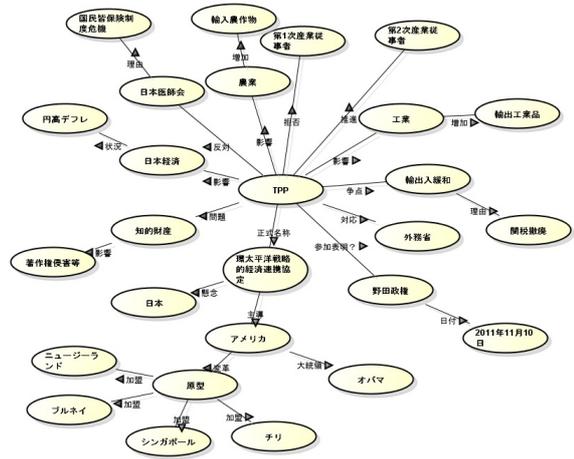
(注1): <http://developer.yahoo.co.jp/webapi/search/>

(注2): <http://mecab.sourceforge.net/>

(注3): <http://d.hatena.ne.jp/tsubosaka/20091223/1261572639>



(a) 検索前コンセプトマップ例



(b) 検索後コンセプトマップ例

図 6 検索でのコンセプトマップの変化の例

4. 提案インタフェースによる利用者実験と評価

実装した検索インタフェースと従来の検索インタフェースを用いて、利用者による比較実験を行う。検索によって得られた情報の整理について、2つのインタフェースでどのような差が出るかを明らかにすることが目的である。本研究では情報を整理する上での表現形式としてコンセプトマップを用いる。課題に沿って検索をしながら、得た情報でコンセプトマップを作成する。課題の検索の前に、課題に対する知識を事前コンセプトマップとして作成し、検索後のコンセプトマップと比較して評価する。また、検索のログについても解析を行い、システム間での違いを明らかにする。本章では実験の内容と結果について述べる。

4.1 コンセプトマップを用いた評価

コンセプトマップとは、ある事柄をノードとし、ノードとノードの間に、関係性を単語、または簡単な文章で付与したリンクを張ったグラフである。図6にコンセプトマップの例を挙げる。図6は実際に実験で利用者が作成したコンセプトマップである。コンセプトマップの作成には *astah* community*^(注4)を用いた。コンセプトマップ作成の手順を容易にするため、リンクの向きについては考慮しないものとして作成をした。ノードやリンクの数など、情報整理の表現形式として定量的な評価を行えるため本研究でコンセプトマップを用いる。評価の仮説としては、提案インタフェースの目的である周辺的な情報の探索が上手く行えるのならば、コンセプトマップは上手く整理されるというものである。具体的にはノードやリンクの増加数を評価する。

4.2 実験概要

2つの検索課題をそれぞれのシステムで行った。利用者には2つの課題Aを提案インタフェース、課題Bを従来インタフェースで行うグループIと、課題Bを提案インタフェース、課題A

を従来インタフェースで行うグループIIの2つに分けた。

表1は実験のパターンを示している。コンセプトマップの作成について、学習によって2つ目の課題で作成するコンセプトマップの方が1つ目のものより広がりやすくなるという明らかな差は見られなかった。そのため、本実験では実験順序を特に考慮しないものとして、IとIIのグループの2つに分ける。この利用者実験では、情報学を専攻する大学生・大学院生の合計15名で、グループIは8名、グループIIは7名で実験を行った。2次元システムは提案インタフェースで、1次元システムは従来インタフェースである。しかし、従来インタフェースではログの取得や処理時間を合わせるため、Yahoo! Web 検索からの検索結果をそのまま提示するインタフェースを実装して用いた。また、キーワードのサジェスト機能については1次元システムでは用いない。

4.2.1 実験課題

実験で行う検索の課題について説明する。検索の課題については、一般的な調べ物についてを対象とする。

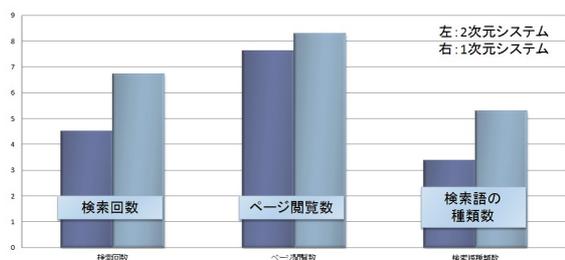
課題Aは経済に関する情報の収集・整理を目的とし、「TPP」についてを調べながらコンセプトマップの作成を行う。

課題Bは環境問題に関する情報の収集・整理を目的とし、「原発問題」についてを調べながらコンセプトマップの作成を行う。それぞれ「TPP」と「原発」という1つのノードから、ノードを追加していき、リンクを作成する。ノードとなる文字列には特に制限はなく、国名やイベント名など、ユーザが関係性を表現できると判断したものを自由に作成してよいと説明を行った。

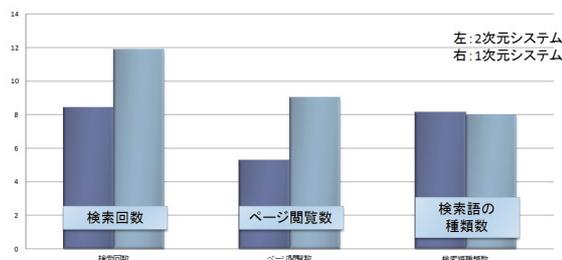
	課題A	課題B
2次元システム	I	II
1次元システム	II	I

表 1 実験パターン

(注4): <http://astah.change-vision.com/ja/product/astah-community.html>

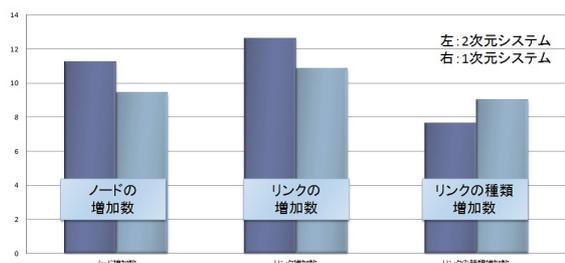


(a) 課題「TPP」での平均グラフ

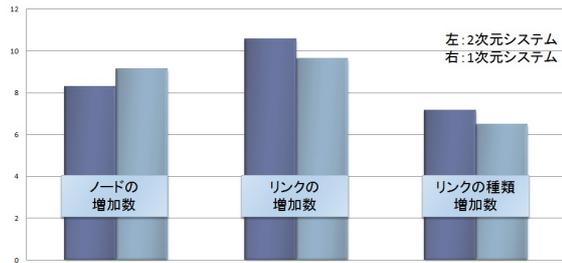


(b) 課題「原発問題」での平均グラフ

図7 検索回数，ページ閲覧数，検索語種類数の利用者平均



(a) 課題「TPP」での平均グラフ



(b) 課題「原発問題」での平均グラフ

図8 ノードの増加数，リンクの増加数，リンクの種類増加数の利用者平均

4.2.2 実験の流れ

たとえばグループIを例としたときの実験の流れとしては、

- (1) 事前アンケートの回答 (10分)
- (2) 2次元システムと1次元システムの説明，および簡単な練習 (20分)
- (3) コンセプトマップの説明，およびツールを使った簡単な練習 (10分)
- (4) 課題Aの流れの説明 (3分)
- (5) 課題Aの事前コンセプトマップの作成 (最大で10分)
- (6) 課題Aの検索とコンセプトマップの作成 (20分)
- (7) 課題Bの流れの説明 (3分)
- (8) 課題Bの事前コンセプトマップの作成 (最大で10分)
- (9) 課題Bの検索とコンセプトマップの作成 (20分)
- (10) 実験後アンケートの回答 (15分)

以上ようになっており，約2時間程度のスケジュールであった。

(5)と(8)については，利用者によって初期状態での課題に対する理解度が異なる可能性を考慮し，検索の開始前にコンセプトマップを作成した。事前コンセプトマップの作成の時間は最大で10分程度であった。(6)と(9)のフェーズでは，事前コンセプトマップから作成を続けてた。また，事前コンセプトマップで作成したノードやリンクの変更は可能であるという条件で行った。

4.3 実験結果と評価

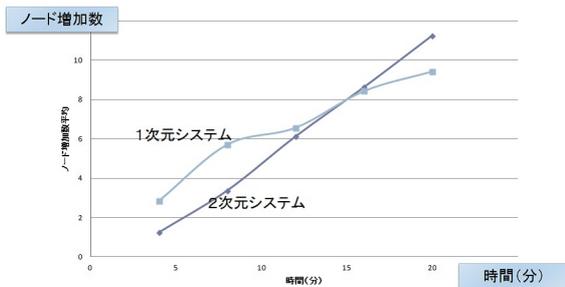
本節では，前節で説明した実験の結果を説明すると共に評価を行う。

4.3.1 検索ログの解析とコンセプトマップの比較結果

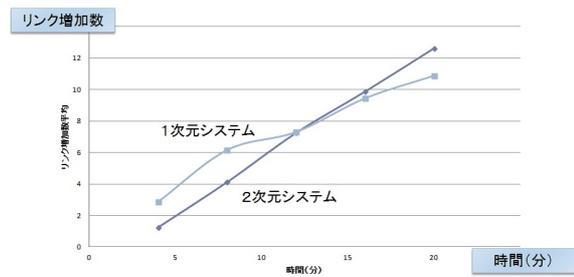
図7は，2次元システムと1次元システムでそれぞれ利用者が検索を実行した回数，リンクを飛んでページを開いた回数，課題内で使われた検索語の種類を棒グラフで表現している。課題「TPP」での2次元システムの利用者は8人で，図7(a)のそれぞれの値は8人の平均である。同様に，1次元システムの利用者は7人で，図7(b)のそれぞれの値は7人の平均である。

図8は，2次元システムと1次元システムで検索を行う前後で増加したノードの数，増加したリンクの数，増加したリンクの種類数を棒グラフで表現している。課題「TPP」での2次元システムの利用者は8人で，図8(a)のそれぞれの値は8人の平均である。同様に，1次元システムの利用者は7人で，図8(b)のそれぞれの値は7人の平均である。リンクの種類数とは，たとえば「TPP」というノードとからリンクを作成するとき「問題」や「影響」といった別々の名前のリンクが多数登場する。一方で「問題」というリンクが多数使われる場合もある。別々の名前のリンクを数え，それをリンクの種類数として図8では表現している。

図7からは，課題「TPP」，課題「原発問題」についても検索の実行回数が少なくなっていることが分かる。ページ閲覧数については，課題「TPP」では両システムで同程度であるが，課題「原発問題」では2次元検索システムの方が少なくなっている。検索語の種類数は，課題「TPP」では2次元検索システムの方が少なく，課題「原発問題」では両システムで同程度となっている。このことから，2次元検索システムでは1次元検索システムに比べて検索を実行する回数が少なくなっていると



(a) 課題「TPP」でのノードの増加数



(b) 課題「TPP」でのリンクの増加数

図9 課題「TPP」での時間によるノード・リンク数の増加

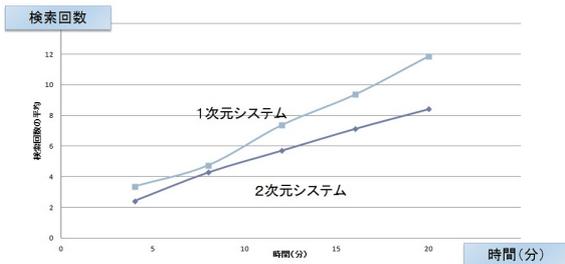


図10 課題「TPP」での時間による検索回数の増加

いえる。また、ページ閲覧数や検索語の種類も少なくなる傾向がある。しかし、t検定を用いて2つのシステムでの有意差を確認したが、検索回数、ページ閲覧数、検索語種類数に有意差は見られなかった。

図8からは、課題「TPP」、課題「原発問題」のどちらでも、グラフに大きな変化がないことが分かる。課題「TPP」については、わずかにノード増加数とリンク増加数に違いが見られるが、t検定によると有意な差ではなかった。

4.3.2 時間経過によるコンセプトマップの変化

図9は、課題「TPP」において、時間の経過によるノードの増加数、リンクの増加数を表現したグラフである。ここでのノード増加数とは、その時間でのノード数から事前コンセプトマップで作成した初期ノードの数である。個人が前提として持っている情報の差ではなく、検索で得た情報を比べるために、増加数を比較した。2次元システムでは8人の平均、1次元システムでは7人の平均である。

課題「原発問題」の時間による増加グラフは両システムでほぼ同じものになり違いが見られなかったが、図9では少しの特徴が見られる。両システム間で有意差のある違いこそは見られなかったが、図9の2つのグラフでは、15分前後までは1次元システムの方がノードとリンク共に2次元システムより増加していることが分かる。しかし、検索終了の段階では2次元システムの方がノードとリンクが増加している。2次元システムは線形的にノードとリンクが増加していくが、1次元システムはノード、リンクの増加が飽和していくことが確認できる。課題「原発問題」ではこのような違いは見られなかった。

課題ごとの違いの原因としては、課題「TPP」が利用者に

とって難しい検索であったことが考えられる。アンケートの結果ではおよそ75パーセントの人が「少し難しかった」「難しかった」と答えている。また、事前コンセプトマップの段階で初期ノードが1つしかないユーザも見られ、「TPP」について全く知らないユーザも存在した。一方で課題「原発問題」については75パーセントの人が「普通だった」から「簡単だった」と答えている。初期ノードの数も課題「原発問題」についての方が多く、課題「TPP」との間に有意差が発生している。このことから『あまり理解していないものに対する検索』、『難しい検索』に対して2次元システムと1次元システムの差が発生すると考えられる。この点については、利用者にアンケートで尋ねた2次元システムの有効性の意見とも一致している。

図10は時間経過による検索回数の増加を表している。2次元システムでは8人の平均、1次元システムでは7人の平均である。12分前後までは有意差・有意傾向のあることを確認した。こちらも課題「TPP」にのみ見られた傾向である。課題が難しかったため、1次元システムでは何回も検索をして情報を探しているが、2次元システムでは可視化された検索結果から情報を探していると考えられる。図7(a)から、1次元検索システムでは検索語の種類が多いこともあり、様々な検索質問を試しながら検索していると推測できる。

5. 考察

2つのシステムで課題を行い、コンセプトマップの比較をしたところ、ノードやリンクの増加量に有意差のある違いは見られなかった。しかし、課題「TPP」ではいくつかの違いの傾向が見られたため、この課題について中心に考察を行う。

図7から有意な差になる可能性がある項目としては、検索の回数と検索語の種類が挙げられる。ノードやリンクの増加を新しい情報の獲得と考えると、2つのシステムではあまり違いが見られなかった。しかし、少ない検索の回数、少ない検索語の種類で同等の情報が得られたと考えれば提案手法が有効である可能性もある。また、今回はノードやリンクの質に注目せず、全てのノードとリンクが等価であるという前提で評価を行なっている。階層関係などに注目し、ノードの質を考慮することも今後の課題といえる。

具体的なマップの広がり方ではなく、別の部分で違いが見ら

れたものもある。2つのシステムで明らかな有意差が出たのは、コンセプトマップの編集回数であった。ノードの追加、リンクの追加、ノード名の変更、リンク名の変更、ノード位置の変更、ノードの削除、リンクの削除の7項目において、課題中に合計で何回の編集を行ったかを確認した。この結果では、圧倒的に2次元システムの方がコンセプトマップの編集回数が多かった。この点を考えると、2次元システムの方が情報を整理しているともいえるが、具体的なノードやリンクの質を評価しなければ有意な情報整理だったのかは判断できない。コンセプトマップを定量的に評価する手法の検討が今後の課題といえる。

6. おわりに

本論文では、検索結果の可視化と対話的な操作により、周辺の情報の発見・整理を支援するインタフェースを提案した。実際に提案インタフェースと従来インタフェースで検索の違いを評価するために、コンセプトマップによる情報の整理を行った。コンセプトマップの広がり方の違いを見ることで、情報の発見と整理にどのような違いがあるかを確認した。

2つの課題AとBを用いて、利用者が作成したコンセプトマップを比較したところ、課題Aでは少しの差が出て、課題Bではほとんど差が出なかった。課題Aは課題Bに比べて難しい検索課題であり、比較的簡単だった課題Bでは検索インタフェースとしての差があまり出なかったと考えられる。課題Aではコンセプトマップのノードの増え方、リンクの増え方の過程に違いが出た。提案インタフェースでは時間に対して一定の伸びがあったが、従来のインタフェースでは時間が経つごとに増加の割合が落ちていくことを確認した。また、提案インタフェースよりも従来インタフェースの方が検索回数も検索語の種類も多くなることも確認した。これにより、少ない検索回数、少ない検索語の種類でも従来インタフェースと同等以上に情報が整理できる可能性を確認した。しかし、統計的に有意といえるほどの差ではなく、今後更なる検討が必要といえる。有意な差としては、課題Aにおけるコンセプトマップの編集回数についてインタフェース間の有意差が生じた。提案インタフェースの方がコンセプトマップの編集回数、特にノードの名前の変更が多かったことを確認した。より情報を整理しているとも考えられるが、コンセプトマップの質を定量的に評価する指標が必要であると考えられる。

検索課題が難しい、または利用者の前提知識が乏しい検索課題ほど従来インタフェースとの差が大きくなることは明らかになっている。今後は更に難しい課題を採用することで、インタフェースの違いを明らかにできると考えられる。また、コンセプトマップの評価方法の更なる検討を行い、分析を行っていく。

文 献

- [1] 林一成, 岩佐英彦, 竹村治雄, 横矢直和. 文書情報の可視化による検索絞り込み支援. 電子情報通信学会技術研究報告, NLC99-80, 99(708), pp.15-20 (2000).
- [2] Roberts, J., Boukhelifa, N., and Rodgers, P.: Multiformal Glyph Based Web Search Result Visualization, the Sixth International Conference on Information Visualisation (IV

- 02), pp. 549-554 IEEE (2002).
- [3] 小林拓海, 佐藤大介, 三末和男, 田中二郎, Web 検索結果の概観提示による情報収集支援インタフェース, 人工知能学会第 19 回全国大会, 3C3-03 (2005).
- [4] 神原啓介, 塚田浩二, XYzon: 商品の比較検討を支援する XY グラフ型検索. インタラクション 2011, 2LNG-4 (2011).
- [5] 吉田大我, 小山聡, 中村聡史, 田中克己, Web 検索結果におけるキーワード出現相関の可視化と対話的な質問変換. 電子情報通信学会第 18 回データ工学ワークショップ (DEWS2007), C7-2 (2007).
- [6] 吉田大我, 小山聡, 中村聡史, 田中克己, Web 検索結果のキーワードマップにおける滑らかな重み付け操作による再ランキング. 情報処理学会第 70 回全国大会, "5-201"-5-202", (2008).
- [7] 中谷文彦, 河合由起子, 熊本忠彦. 柔軟な Web コンテンツ検索のためのレーダーチャート検索システムの提案. 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008), B5-4 (2008).
- [8] Yuka Egusa, Hitomi Saito, Masao Takaku, Hitoshi Terai, Makiko Miwa, Noriko Kando, Using a Concept Map to Evaluate Exploratory Search. Proceedings of the Third Symposium on Information Interaction in Context (IiIX 2010), pp.175-184 (2010).
- [9] 江草由佳, 齋藤ひとみ, 中島諒, 高久雅生, 神門典子, 三輪眞木子, Web 情報探索前後のコンセプトマップの分析:探索者によるノード同定結果を用いて. 研究報告情報基礎とアクセス技術 (IFAT), 2011-IFAT-104, pp.1-7, (2011).