

多様な検索結果を表示する可視化手法に関する一検討

玉田 雄基^{†,††} 佐藤 哲司^{††}

† 筑波大学情報学群知識情報・図書館学類 〒305-0550 茨城県つくば市春日1番2号

†† 筑波大学 図書館情報メディア系 〒305-0550 茨城県つくば市春日1番2号

E-mail: †{tamada,satoh}@ce.slis.tsukuba.ac.jp

あらまし ユーザが求める多様で複雑な情報要求に応えるには、単純なキーワード検索で特定の Web ページを探すだけでは不十分であり、関連するいくつかの Web ページを辿りながら横断的に閲覧する探索的な情報検索を行わなければならない。複数ページにまたがる情報を閲覧・収集するためには、ハイパーリンクを辿る方法などがある。しかし、ハイパーリンクを辿って閲覧ページを次々に切り替えることは、情報の整理や理解が難しくなるなどの問題がある。本研究では、検索条件を満たす候補ページへのリンク集である検索結果の横断的な検索の支援を目的に、具体的には、検索結果を俯瞰しながら簡単に情報の比較・理解が出来る、新しい検索インターフェースを提案する。検索結果を二次元平面に配置し、ハイパーリンク関係に基づいて個々のページ間の関係や全体の分布を可視化する。

キーワード 情報の可視化, 情報検索, インタラクション

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には膨大な数の Web 文書が存在し、文書数は今なお爆発的に増え続けている。このような状況の中、検索エンジンを使用するユーザは、自分の求めている情報を得るために検索結果の中から所望の情報を取捨選択をしなければならない。また、ユーザが検索で求める情報は多様化・複雑化してきており、単一の Web ページから情報を得るだけでは、ユーザの要求を完全に満たすことは難しい。このため、ユーザは 1 つの Web ページだけではなく、複数の Web ページを横断的に閲覧し、必要な情報を選択していくことになる。例えば、旅行プランを立案するには、行き先やホテル、交通手段の選択など、様々な関連する情報を相互に見比べながら取捨選択しなければならない。

煩雑な情報が混在している Web 情報空間の中で、比較的整理された情報源として Wikipedia がある。Wikipedia はウィキペディア財団が主催し、すべてのユーザが無料で自由に編集・閲覧できるインターネット百科事典である。現在ではインターネット上から、なんらかの知識を得る際に欠かすことのできないサービスとなっている。Wikipedia においても、横断的な閲覧を可能とするハイパーリンクが用意されている。それを辿っていくことで横断的に情報を取得し、情報要求を満たしていくことが出来る。

ハイパーリンクを辿る操作は、閲覧ページを次々に切り替えることである。ハイパーリンクを辿る過程では、閲覧しかかりのページが増加し、情報の整理を難しくする。このように、複数ページを同時に閲覧しなければならない場合も多く、複数ページを同時に閲覧するための支援が重要となる。

以上のことから、検索エンジンから大量の検索結果が提示された場合でも、何度も検索行動を繰り返すことなく、情報を的確に把握・整理を行い、情報要求を満たしていくことが出来るインターフェースが求められている。商用検索エンジンの検索

結果の提示法は、検索結果を上から順番に一次元のリスト構造で提示し、その提示順に各文書間の関係性はない。このため、様々な検索結果や始めの検索結果の周辺情報から、情報を断片的に取得し、情報要求を満たしていく検索行動は難しいと言える。

本研究では、検索条件を満たす候補ページへのリンク集である検索結果を横断的に閲覧することを支援し、さらには連続する検索行動で得られる結果集合の間の関係を整理することを目的として、検索結果を俯瞰しながら簡単に情報の比較・理解が出来る、新しい検索インターフェースを提案する。検索結果を二次元平面に配置し、ハイパーリンク関係に基づいて個々のページ間の関係や全体の分布を可視化する。これまで一次元表示であった検索結果を二次元平面に可視化することで情報間の関連を分かりやすくし、可視化したページを見ながら閲覧するページを選択可能とすることで、情報の整理が容易にする。

実現したプロトタイプシステムでは、検索支援となる情報源に Wikipedia^(注1) を使用し、検索結果をページアイコンとして可視化する。そしてユーザのインタラクションによって閲覧するページを次々に変更ながら、情報要求を満たしていく。また、閲覧して気に入ったページを保持しておくことで情報の比較・検討も行うことが出来る。これにより、情報を横断的に検索し、ユーザの情報選択を助けるインタラクティブな検索を実現する。

本論文では 2 章で関連研究について述べ、本研究の位置づけを明らかにする。3 章では提案手法について述べる。4 章では提案システムの実装、システムを用いた実験について示す。5 章では結果を考察し、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

本研究では、検索結果のページをそれぞれアイコン化して二

(注1) : <http://ja.wikipedia.org/>

次元平面に提示するが、検索結果を可視化する研究は多数存在している。

林ら [1] は、文書情報の可視化による検索結果の絞込みを支援する手法を提案している。正解文書中に出現する単語に類似する単語を列、各文書を行としたマトリクス形式で検索結果を提示する。ユーザは、提示された検索結果の中からある文書を適合文書として指定することで、それに類似する単語を持つ文書を並べ替えて提示する。また、ある単語を追加して検索することで得られる検索件数を求め、絞込み率を計算する。検索結果中での単語の出現頻度と、絞り込み率の二つの軸によって二次元平面を構成し、単語を配置する。これにより、絞り込みのための単語を視覚的な情報で与えることを目指している。

山下ら [2] は、文書間のリンク構造と、文書中に出現する用語の共起関係に着目し、結びつきの強いサイトを二次元平面上に視覚化することで検索支援を行う手法を提案している。この手法は、同じハイパーリンクでもアンカーサイトと関連の強いサイトへのリンクと、ポータルなどへのリンクでは重みが違うと考え、リンク構造と共起関係を平面上に視覚的に表示することで、情報検索支援を行っている。

早風ら [3] は、Web サイトに張られたリンクにより Web サイト間の関連度の計算を行い、基準となる Web サイトと関連性の高いサイトを可視化する手法を提案している。Web サイト間の関連度は、リンクを Web サイト単位の有向グラフとして扱うことで到達ステップ数から導き出し、その関連度の強さによって Web サイトのタイトルバーの色、前後関係を変化させている。ユーザは指定した Web サイトと関連性の高い Web サイトの検索が容易となり、サイト間の相互関係を直観的に理解することが出来る。

砂山ら [4] は、ある用語について調べている場合、その意味だけを知りたい場合、または意味を含めた背景情報まで知りたい場合とでは、必要とする情報源に違いが出てくると考え、選択中の Web ページのリンク先の情報がどのような記述になっているのを知ることが必要であると考えた。そのため、Web ページ間の相対的な具体抽象関係を定義し、検索語について集められた Web ページを、ユーザが着目する Web ページにと比較して、「具体的な Web ページ」と「抽象的な Web ページ」の評価を与え、視覚化するシステムを考案している。ユーザは「具体的」または「抽象的」という情報をから自分の要求にあう Web ページの内容を確認していくことができ、情報収集の効率化が期待出来る。

松尾ら [5] は、ユーザが検索に使用する質問キーワードは、ユーザの興味は反映しているものの、多様なページが混在している Web 空間において必ずしもユーザが求める情報を正確に記述しているとは限らないとする考えに基づいて、検索結果の視覚化を行なっている。そのため、質問キーワードで得られる情報と、その周辺情報も一緒に視覚表示する。それにより、ユーザが求めている情報の発見、またユーザ本来の目的により近い情報を検索するための質問の生成を支援する。他の研究と異なる点は、この「質問を生成する支援を行なう」という点にある。ユーザは、視覚表示されたページ集合を閲覧しながら、情報要

求を満たすことが出来る他のキーワードを発見しながら質問を修正し、漸進的に検索を進めていくことが出来る。これは、検索結果可視化の研究でありながら、2.2 節で説明する、「インタラクティブな操作」にも関係がある研究と言える。

2.2 インタラクティブな操作

本研究では、検索結果ページを次々に閲覧するインタラクティブな操作を行うことで、検索結果の表示・変化を行っていく。このようなインタラクティブな操作によって結果を変化させていく手法も多く提案されている。

塩澤ら [6] は、インタラクティブに食材の優先度を調整することによって、料理レシピの検索結果を見渡しながら、ユーザが作りたいレシピを選べる検索手法を提案している。この手法は、食材の名称や種類からレシピを検索するという従来の手法に加えて、ユーザは入力するクエリに“優先度”を付与して検索を行うというものである。優先度には +10 から -10 までの範囲が設定しており、入力されたクエリの優先度を考慮した検索結果が表示される。そしてユーザは、優先度やクエリをインタラクティブに操作することで可視化された検索結果から所望のレシピを選択出来る。

軽部ら [7] は、検索対象分野の知識が浅いユーザを支援する検索支援手法を提案している。この手法では、情報要求に適合/非適合した文書片をユーザが検索結果のページ内からインタラクティブに取捨選択することにより、動的に検索結果をリランキングする。ユーザは、検索結果から得られる適合文書片の追加、削除をインタラクティブに同一画面内で行うことで検索結果を確認していき、徐々に情報要求を満たしていくことが可能である。

野間口ら [8] は、レシピ検索をする中で、「冷蔵庫の残り物を使える料理」などの漠然としている情報探索では、検索結果を一次元リスト構造で表すよりも、比較・検討が容易である、二次元に視覚表示する手法が適切であると考え、情報視覚化を利用したレシピ探索支援のためのシステムを考案している。このシステムでは、検索過程を、ユーザからの検索要求である個々の食材と、検索結果であるレシピをノードとするグラフとして視覚化する。ユーザは、思いついたクエリを簡便にノードとして追加することで検索を行なえる。その後、ユーザは求めるレシピにたどり着くまで、視覚化された結果をインタラクティブに操作し、比較・検討することが可能である。

2.3 本研究の位置づけ

本研究では、可視化した検索結果を見ながら、ユーザが自身の情報要求に合わせて閲覧ページをインタラクティブに変更させていくことが出来る、検索結果の可視化手法を提案する。可視化表示を行うことで情報の一覧性の向上、検索結果同士の関係性を知り、インタラクティブな操作を行うことによって情報の整理を行う。

また、本研究と塩澤ら、軽部らの研究は、検索結果へのインタラクションによって情報要求を満たしていくという点で類似しているが、可視化表示された検索結果を直接操作することで検索結果の変更、また取得した情報の比較を容易にするという点で本研究は異なる。また、野間口らの研究は可視化後に結果

の操作を行なっているなど本研究と狙いは類似しているが、レシビ検索は単一の検索結果を探して検索を行なっているのに対し、本研究は横断的な検索を対象としている。

以上述べたように、本研究は、視覚的に検索結果を表現し、それらを俯瞰しながら閲覧ページを動的に変化・比較を行い、情報要求を満たしていくという点に特徴がある。

3. 提案手法

3.1 情報検索モデル

本研究では、Wikipedia の検索結果ページを二次元平面に可視化表示し、インタラクションによって表示結果を動的に変更する。従来の検索方法では、検索結果ページを閲覧し、情報要求を満たしていくことが一般的である。しかし、検索結果が自分の要求と合致しなかった場合、大部分のユーザは、ハイパーリンクを辿って周辺情報を横断的に閲覧する、検索語を変更して再検索を行なう。そして、ハイパーリンクを辿っていった場合、ページを辿っていくたびに閲覧ページが切り替わってしまうため、情報の比較や整理が困難になる。そこで、検索結果を多次元空間に表現すれば、視覚的に情報を把握することが可能となり、必要とする情報を分析、選択が容易になるのではないかと考えた。しかし、三次元空間やそれ以上の次元を表現することは、逆にユーザにとって複雑すぎて扱いづらい可能性がある。また、インタラクションによってレイアウトを動的に変化させることを考えた時、操作自体も複雑化すると考えられる。ユーザに容易に、分かりやすく情報を提供することが出来るのは、二次元平面であると考え、ページを二次元に可視化する手法を考えた。

その手法の一例として、図 1 がある。この図は適合文書を検索するため、ハイパーリンクを回遊しているときの検索結果閲覧方法の違いを比較したものである。従来の方法では、上述したとおり、リンクを辿ることで何度も画面遷移を繰り返して適合文書を検索している。それに対して提案手法では、各ページをハイパーリンク関係によって繋ぎ、可視化する。そしてページ間の関係を空間的に把握、整理しながらブラウジングを行うことが出来る。そのため、リンクを辿らなくても文書概略を把握することができ、適合文書を容易に捜し出すことが出来る。

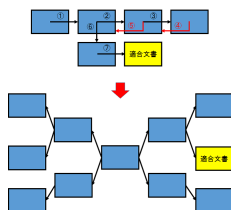


図 1 検索結果閲覧方法

従来の検索手法のフローチャートと、本研究の検索手法のフローチャートを図 2、図 3 で比較する。従来の手法では検索語を入力した後、検索結果を閲覧する。そして、情報要求を満たすまでハイパーリンクを辿っていく。その検索行動の中で、ハイパーリンクを辿る場合も、検索結果を変更する場合も、一度その行動を行って見なければどのような情報があるのか知ることとは出来ない。そのため、検索しては元の画面に戻って、またほかの検索を行うという状況が見うけられる。それに対して提案手法では、同一画面内で可視化結果の選択、閲覧が可能であるため、ユーザは内容を類推しながら、次々に連続してページを閲覧することが出来る。

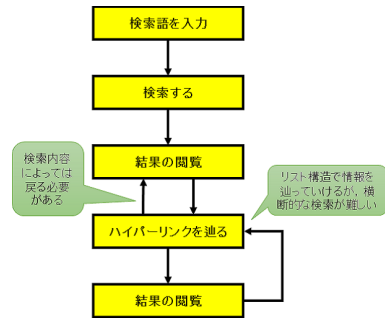


図 2 従来の検索フローチャート

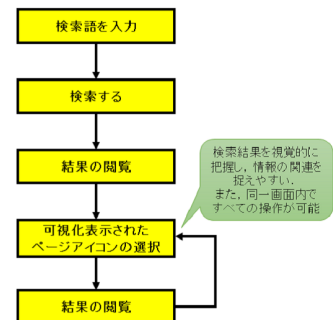


図 3 提案手法の検索フローチャート

3.2 検索結果の可視化

検索結果の関係性を表すために、ハイパーリンクを使用して、検索結果を一次元リスト構造ではなく、二次元に可視化する。ハイパーリンク構造を持つ情報源をノードでつなぎ、それらをアイコン化し、可視化表示することで検索結果の可視化を実現する。また、ノードでつながれたアイコンを俯瞰しながらページを閲覧することで、情報間の関係を理解しながら検索行動を進めていくことが出来る。

具体的な実現方法としては、Wikipedia のリンク関係とタイトル、本文が格納された DB を作成し、ユーザが入力した検索語から作成した DB にアクセスし、関連のあるページタイトルを取得する。そして、取得したキーワードからリンク関係のあるキーワードを DB から取得し、arbor.js という javascript のライブラリを使用して可視化を実現する。

ここで、一般的に言われている、情報を可視化することで得

られるメリットについて述べる。人間の知覚というものはイメージに対してとても敏感であり、高い順応性を示している。そのため、視覚的な表現はテキストに比べて様々な情報を効率よく伝達することが出来る。例えば、棒グラフや折れ線グラフは、数字の対応表に比べ、潜在的なデータへの気づきを喚起する。これはほんの一例ではあるが、テキストデータのみを見て記憶を行なうより、テキストと視覚情報をあわせて人の記憶へ影響を及ぼすことが出来る検索を目指し、本研究は可視化という手法を選択した。

3.3 提案システムの機能

提案システムにおける初期状態を図4示す。ここで言う初期状態とは、パソコンを起動してから、初めてシステムをスタートさせた時の状態のことであり、まだ検索は行われていない。左上には検索語入力フォームが配置され、ユーザはここに検索語を入力する。左下に置いてあるウィンドウには可視化結果が表示され、ユーザはこのウィンドウを見て、操作を行いながら検索を進めていく。この検索結果を可視化した結果を表示する部分を、関連性表示部と名づける。

右下に置いてあるウィンドウは選択されている検索結果ページを閲覧する場所となっており、本研究ではこの領域を詳細ページ表示部と名づける。

最後に右上の領域にある複数のウィンドウはページ保存部名づけ、ここはユーザが検索行動を繰り返す中で、気に入ったページを保存する場所とである。

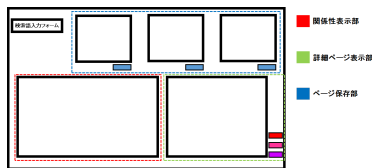


図4 初期ウィンドウ

3.3.1 関連性表示部

関連性表示部に表示される具体例を図5、図6に示す。

図5では地理学を検索語として、関連のある検索語を可視化した結果を示す。“地理学”と表示されたアイコンを中心として放射状にリンク関係が可視化されている。

それに対して、図6では地理学を検索語として、自分から二つ離れたところまでのリンク関係を可視化した結果を表示している。本研究では、検索語を赤いアイコンで表示し、第一階層と名づける。その後ハイパーリンクの構造が第二階層、第三階層と深くなっていくたびに、緑色のアイコン、青色のアイコンと色彩が変化していく。また、ハイパーリンク構造の第二階層に存在するキーワードが、第三階層にも表示される場合は、新たに同じキーワードのアイコンを複製することはせず、リンク関係があることを可視的に示しているエッジだけを引くことにする。

図5と図6を見比べて、地理学に関係のあるページを列挙するよりも、リンク構造が3階層程度存在するほうが、ページ間の関係が見えてくる。そのため本研究では、図6のように、ハ

イパーリンク関係を三階層まで可視化したものを用いることとする。

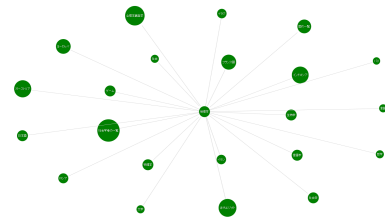


図5 二階層ハイパーリンク構造

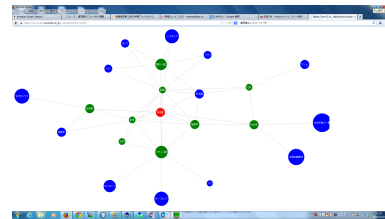


図6 三階層ハイパーリンク構造

3.3.2 詳細ページ表示部

詳細ページ表示部には、ユーザが関連性表示部において発見し、選択したキーワードの wikipedia ページが表示される。ユーザは関連性表示部に隣接するこの領域の表示ページを次々に変化させながら検索行動を繰り返し、自らの情報要求を満たしていく。

具体的な実現方法としては、DB から取得したハイパーリンク関係があるキーワード群に動的に `キーワード` というタグを付与する。これにより、キーワードをクリックしたときにそのキーワードの wikipedia ページを詳細ページ表示部に表示する。

3.3.3 ページ保存部

ページ保存部は、ユーザが検索行動を繰り返す中で、気に入ったページを保存しておく領域となっている。詳細ページ表示部の横にある保存ボタンをクリックすることで、詳細表示されている wikipedia ページのキーワードとページ概要が保存される。また保存されたページを再び詳細に閲覧したい場合は、ウィンドウの下についているボタンをクリックすることで、保存ページの情報を詳細表示部に送信する。

詳細ページ表示部からページ保存部にページを移動させる時は、javascript のイベントアクションでボタンがクリックされた時に、ページタイトルと概要を抽出してウィンドウに書き出す。ページ保存部から詳細ページ表示部にページを移動させる時は、ウィンドウの下に配置されているボタンに wikipedia ページのリンクを動的に付与し、詳細ページ表示部のウィンドウ名をターゲットに指定しておく。そして、同様にイベントアクションで動作するようにしておくことで実現する。

ページがどのウィンドウに移動するのかをイメージ化したものを図7に示す。wikipedia ページはボタンの色と同じ色のウィンドウに表示される。

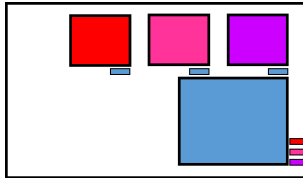


図 7 ページの保存

3.4 インタラクティブな操作

3.4.1 インタラクティブな操作の概要

従来の検索システムでは、自分が入力したクエリによって応答(検索結果)を返す。そしてユーザは検索エンジンから戻ってきた膨大な検索結果の中から、情報を取捨選択し、自身の情報要求を満たしていく。この行動は何の変哲もない、普通の行動ではあるが、思いのほかユーザに付加がかかるものである。例えば検索対象の知識が拙く、範囲が幅広い場合などは、ひとつの情報源だけではなく、その周辺情報を検索することで、検索対象をより詳しく知ることが出来る。しかし、従来の検索システムは検索結果をユーザに送信し、ユーザは、その中から選択を迫られるという一方向な検索になっている。

本研究では、ユーザがストレスなく自身の情報要求を満たすことが出来るよう、検索結果をシームレスに閲覧出来るインタラクティブなシステムを提案する。本研究で言うインタラクティブな操作とは、検索結果を可視化したものを直接操作して閲覧ページを変更させること。その操作に付随して起こる、検索結果のシームレスな変更のことを言う。また、インタラクティブな操作とは別に、過去に閲覧していたページを取り置いておくことで、情報の把握、さらには比較・検討を容易に行うことが出来る機能も追加している。これにより、ユーザは一方向に送りつけられる検索結果の中からではなく、ユーザの操作によって応答を変化させることが出来る、対話的で双方向な検索システムになっている。

3.4.2 インタラクティブな画面遷移

検索語を入力し、検索を行った場合の画面遷移を図 8 に示す。今回の例では「文書 1」を最もよく表した検索語を入力したこととする。これにより、関連性表示部には、「文書 1」をスタートとしたハイパーリンク関係によって可視化された検索結果キーワード群が表示される。また、現在選択されている「文書 1」の内容は、詳細ページ表示部に表示され、マウスでスクロールすることで通常通り閲覧可能である。

可視化結果は 3.2 章で記述した通り、「文書 1」を中心としてハイパーリンク関係を放射状に表現し、各ページを線で繋ぐ。また、図中で表現されているアイコンには、説明のために番号を付与してあるが、実際のシステム中では、ページタイトルがラベルとして付与されている。

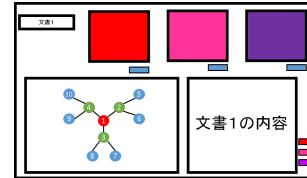


図 8 画面遷移 1

選択キーワード変更させた時の例を図 9 に示す。この図では、図 8 で選択されていた「文書 1」から、ひとつ先の階層にある「文書 3」を選択したときの画面遷移を示している。「文書 1」から「文書 3」に選択ページを変更することで、詳細ページ表示部には文書 3 の内容が表示される。

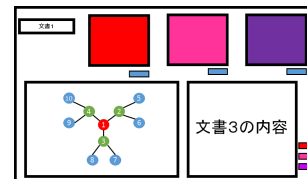


図 9 画面遷移 2

詳細ページ表示部に表示されているページを保存する時の例を図 10 に示す。この図では、ユーザが「文書 3」の内容をページ保存部に取り置いた時の画面遷移を示している。詳細ページ表示部の隣に 3 つの保存ボタンがあり、保存ボタンをクリックするとページ保存部のウィンドウにページタイトルとページの概要を保存する。

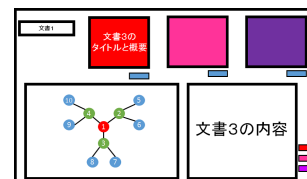


図 10 画面遷移 3

4. 実装

本章では、提案手法を実装したプロトタイプシステムによる評価について述べる。4.1 節で使用するデータセットについて述べ、4.2 節で提案システムによる実装の結果について述べる。4.3 節で提案システムの有用性の検証を行なう。

4.1 使用するデータセット

本研究では、検索対象とするデータセットに Wikipedia のダンプデータを使用する。Wikipedia のデータは、2013 年 10 月 23 日の 00:28:26 の時点の最新版を使用し、記事総数は 1,757,201 件である。データセットの構成内容を表 4.2 から表 4.4 に示す。

page, revision, old の 3 つのテーブルからデータセットは構成されており、ページタイトルと本文が直接紐付けられていない。「ページタイトル → page_id → rev_id → 記事本文」と辿ることを避けるため、MySQL の機能の一つであるビューを使用して、ページタイトルから直接に記事本文にアクセスできるようにした。このビューのスキーマを表 5 に示す。

表 1 テーブル : page の構成

フィールド名	フィールドの詳細
page_ id	記事をユニークに識別する主キー
page_ namespace	ページの名前空間の番号を格納する
page_ title	ページのタイトル
page_ restrictions	編集等の資格を表すパーミッションキーのセット
page_ counter	ページが閲覧された回数
page_ is_ redirect	この記事がリダイレクトであるか否か
page_ is_ new	ページが新規のエントリか否かを示す
page_ random	ランダムな 0 と 1 の間の小数值
page_ touched	タイムスタンプ
page_ latest	rev_ id への外部キー
page_ len	ソーステキストの圧縮していないバイト単位での長さ

表 2 テーブル : revision の構成

フィールド名	フィールドの詳細
rev_ id	リビジョンのために主キーを保持
rev_ page	リビジョンに関係する page テーブル の参照を保持
rev_ comment	編集者の 編集内容の要約 (編集者のリビジョンでのコメント) を保持
rev_ text_ id	Text テーブルの中の old_ id へのポインタ
rev_ user	編集をした利用者の user_ id と同じ
rev_ user_ text	編集者の IP アドレスを保持
rev_ timestamp	編集の タイムスタンプ を保持
rev_ minor_ edit	‘これは細部の編集です’ チェックボックスにマークをしたかどうかを記録
rev_ deleted	RevisionDelete システムに予約されているフィールド
rev_ len	リビジョンの長さを、バイト単位で格納
rev_ parent_ id	ツリー構造 (The Adjacency List Model) のサポートを追加するために使用

表 3 テーブル : old の構成

フィールド名	フィールドの詳細
old_ id	主キー
old_ text	ページの Wiki テキスト
old_ flags	フラグをカンマ区切りで列挙したもの

表 4 テーブル : pagelinks の構成

フィールド名	フィールドの詳細
pl_ from	記事をユニークに識別する主キー
pl_ namespace	名前領域
pl_ title	記事にリンクされているページのページタイトル

表 5 作成したビューの構成

フィールド名	フィールドの詳細
page_ title	ページのタイトル
old_ text	ページの Wiki テキスト

4.2 プロトタイプシステムの実装

実際に作成したプロトタイプシステムの構成図を図 11 に示す。ユーザが実際に入力や結果の出力、インタラクティブな操作を行なうプログラムがブラウザ上でインターフェースとして提供され、処理を行なうシステムがサーバ上にある。まず、

ユーザから入力された検索語を、ブラウザの Javascript を介してサーバサイドシステムに渡す。サーバサイドで動いている PHP プログラムは、検索語を受け取り、MySQL を介して DB から、検索語に関係するページ群を取得する。そしてそれらのページ群のハイパーリンク関係を計算し、クライアントサイドに送り返す。それを受け取ったブラウザ上のプログラムが実際にページ関係の可視化を行なう。以上のように、ユーザは検索語を入力することで、二次元平面に可視化された検索結果を得ることが出来る。

次に可視化結果を操作したときの動作について説明する。元々選択されていたページを取り替え、ほかのページを選択した場合、詳細ページ表示部の表示が変化する。また、ユーザが検索を繰り返す中で発見した、お気に入りページをページ保存部に保存しておくことが出来る。ユーザはこの機能を使用して、情報の比較、検討を行なうことが出来る。

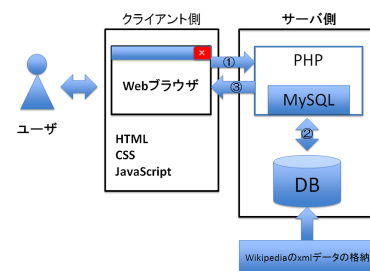


図 11 提案システムの構成図

検索語入力フォームに「数学」と検索語を与えて、検索を行なった結果が図 12 である。左下に配置してある関連性表示部には、赤いアイコンで示された「数学」を中心にして、ハイパーリンク関係があるキーワードのアイコンが放射状に配置されている。また、可視化された検索結果だけではなく、列挙した関連キーワードからでも情報を見つけ出すのを助けることが出来ると考え、プロトタイプの実装では関連キーワードを列挙した関連キーワード BOX を付け加えた。

右下に配置してある詳細ページ表示部には、「数学」の wikipedia ページが表示されていることが分かる。これはスクロールによって全体を閲覧することが出来るようになっている。

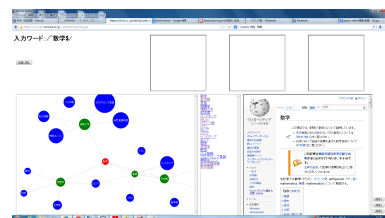


図 12 「数学」と検索した時のウィンドウ

閲覧するページを「数学」から「生物学」に変化させたものが図 13 である。本来は、「数学」の wikipedia ページから「生物学」の wikipedia ページに到達には、「数学」→「物理学」→

「生物学」とリンクを辿っていく必要がある。しかしこのプロトタイプシステムでは、二つ先のリンク関係まで可視化して選択出来るため、「数学」から直接「生物学」に到達することが出来る。図中のアイコンは、入力した検索語で検索された結果を赤色、その検索結果（ページ）から直接リンクがあるページを緑色、そして「生物学」のように検索語から二つ先のリンク関係を青色のアイコンで表している。

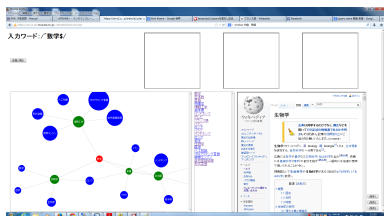


図 13 「生物学」に閲覧ページを変更した時のウィンドウ

詳細ページ表示部において、閲覧したページが検索要求に合致あるいは、極めて合致していると思われるページが提示されたら、そのページを縮小したイメージをページ保存部に取り置きする。現在、ページ保存部に保存できるページ数は3としていたので、3ページまではそのまま保存できる。既に保存してある3つのページよりも検索要求に合致したページを閲覧した場合は、既に保存してあるページと入れ替えて、新しいページを保存する。

4.3 提案法の有用性

4.3.1 関連性に基づく情報探索

情報探索では、関連（ハイパーリンク）を次々に辿っていくインタラクティブな操作が必要となる。例えば、「地理学」と検索語したときの関連性には、図14のようになる。従来の情報検索で「地理学」と検索した場合には、国やそれらに関係する学問の情報が多くでてくる。しかし、本研究では、青色アイコンで示された、ゲームのようなキーワードも散見され、思いがけない関連を見つけ出すことが出来るであろう。

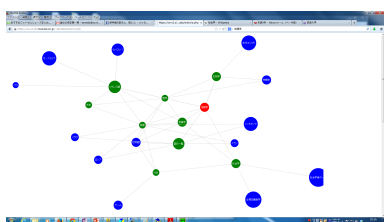


図 14 地理学の関連性

4.3.2 網羅性が必要となる情報探索

提案システムでは、2つ先のノード（見出し語）までを一覧することが出来る。このことは、ある事柄に関連する情報を網羅的に探す際に、提案システムは有効に機能すると言える。例えば、「著作権の保護期間」を検索語したときの関連性には、図15のようになる。この場合には、「著作権の保護期間」と直接ハイパーリンク関係がある「映画」や「テレビ」だけでなく、二つ先のリンクである「音楽」や「文学」などに着眼し、従来の

情報検索のより幅広い「著作権の保護期間」を知ることが出来るはずである。

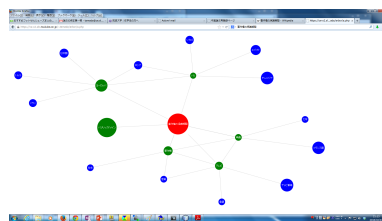


図 15 著作権の保護期間の関連性

5. 考察

本研究では、検索結果をハイパーリンク構造によって可視化表示、その結果を直接操作することで、周辺的な情報の発見・情報間の関係の把握を支援するインターフェースを提案した。

5.1 検索結果の可視化

図3.4と図3.5、また図4.3と図4.4に示したように、検索結果可視化には成功し、従来の検索では1stepでは到達することが出来ない情報をユーザに提供することが出来た。これによりユーザは情報間の関係性を把握し、情報の取捨選択が容易に出来るようになった。また、現在はリンク関係を検索語から二つ先のキーワードまでにサーバサイドで設定しているが、ユーザがそれ以上先まで知りたい場合には、設定を変更することも可能である。

検索結果の可視化に関する課題は二つある。一つはリンク関係があるキーワードの取得方法を再考する必要があること。現在はページ中に貼られているハイパーリンク数をDB内でカウントし、その数が大きいキーワードから取得、表示を行なっている。そのため、人名や商品などの前に多く使用される地名や国名が大きなスコアになってしまい、提案システムで検索を行なうと、それらが頻出してしまいうという問題があった。

二つ目の課題としては、システムのユーザエクスペリエンスの問題である。林ら[9]は、検索結果を可視化し、情報を整理するためにそれらに対話的に操作出来る、検索インターフェースを構築した。しかし、そのシステムの使用後に調査されたアンケートでは、従来の一次元検索システムのほうが使いやすい、という意見もいくつか存在していた。その理由としては、やはりgoogleやYahoo!といった大規模検索エンジンがすべて一次元リスト構造であり、「慣れているから」という点が大きかった。そのため、今後更なるユーザエクスペリエンス向上を目指すためには、この点に留意する必要があるだろう。

5.2 インタラクティブな操作

インタラクティブな操作として提案していた詳細ページ表示部には、図4.3と図4.4のように、wikipediaのページを表示させることが可能となっている。ユーザはこの機能を使うことで、同一画面上で繰り返しページを閲覧することが可能となった。

インタラクティブな操作に関する課題としては、ページ保存部の実装を完了させ、同一ウィンドウ内でシームレスにwikipediaページをやり取りすることである。この機能によってユーザは情報の比較・検討を行なうため、早急に実装するが必要である。

6. ま と め

本研究は, 検索条件を満たす候補ページへのリンク集である検索結果の横断的検索の支援, さらには連続する検索行動で得られる結果集合の間の関係を整理することを目的として, 検索結果を俯瞰しながら簡単に情報の比較・理解が出来る, 新しい検索インターフェースを提案し, システムの構築した. 検索結果のページをアイコン化して可視化表示し, それらの関係性を把握しながら, 表示ページを次々と取り替えながら検索を進めることで, 横断的に検索をすることが出来ることを確認した.

具体的には検索結果の可視化手法と, インタラクティブな操作を実現するページの表示・保存手法の提案を行なった. 検索結果の可視化手法としては, 検索語のリンク関係があるキーワードを DB から抽出し, それらをリンク構造の階層によって色分けして可視化した. これにより一覧しただけでリンク構造の把握が容易となり, 閲覧する情報の選択を助ける効果をねらった.

提案するインターフェースのプロトタイプを実装し, 動作の実験をした. 具体的には, 検索結果が可視化表示されたかを確認した. その結果, 検索語から二つ先のリンク関係があるキーワードを取得し, キーワードのラベルが付与されたアイコンを可視化表示することが出来た. また, 選択キーワードを変更することで, 閲覧ページをシームレスに変化させることができ, 横断的検索を支援出来ることを確認した.

今後の課題としては, ページ保存部の実装, キーワード取得方法の変更, ならびに定量的評価を行なうことである.

謝 辞

本研究の一部は, JSPS 科研費 25280110 の助成を受けたものである.

文 献

- [1] 林一成, 岩佐英彦, 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 電子情報通信学会技術研究報告 NLC99-80,99(708), pp. 15-20, 2008.
- [2] 山下仁志, 石田一成, 岡田勇, 太田敏澄. リンク構造と共起関係を用いた web 空間の視覚化. 情報処理学会 情報処理学会研究報告 情報学基礎研究会報告 2004(36), pp. 95-101, 2004.
- [3] 早風一志, 奥平雅士. Web サイト間の関連度可視化に関する一検討. 映像情報メディア学会 映像情報メディア学会技術報告 33(11), pp. 53-56, 2009.
- [4] 砂山渡, 鯨島聡志, 西原陽子. Web ページ間の相対的な具体抽象関係の視覚化による情報収集支援. 電子情報通信学会 電子情報通信学会論文誌 D 情報・システム J92-D(3), pp. 271-280, 2009.
- [5] 松生泰典, 是津耕司, 小山聡, 田中克己. Web サイト間の関連度可視化に関する一検討. 情報処理学会 情報処理学会研究報告 データベース・システム研究報告会 2004(71), pp. 53-56, 2009.
- [6] 塩澤秀和. 食材の優先度を考慮したビジュアルな料理レシピ検索インタフェース. 電子情報通信学会 電子情報学会論文誌 A 基礎・境界 J94-A(7), pp. 458-466, 2011.
- [7] 軽部孝典, 志築文太郎, 田中二郎. 検索結果のインタラクティブな評価に基づくランキングインターフェース. 情報処理学会 情報処理学会第 70 回全国大会 "1-489"- "1-490", 2008.
- [8] 野間口佑也, 星野准一. Graphicalrecipes: レシピ探索支援のための視覚化システム. 学術科学会論文誌 Vol.7(2008) No.2, pp. 43-54, 2008.
- [9] 林大策, 福原知宏, 佐藤哲司. Web 検索結果とその周辺情報の近傍的内包表現とその視覚化. 情報処理学会 マルチメディア, 分