情 報 学 基 礎　59－3
デジタルドキュメント　24－3
（2000．7．28）

ハイパーリンクとアンカーテキストを利用した情報検索とランキングの一手法

風間 一洋，　原田 昌紀，　佐藤 進也
NTT 未来ねっと研究所

本稿では，ハイパーリンクとそれに関連づけられたアンカーテキストを使って検索することで，ランキングの質を向上する手法について述べる．さらに，サーチエンジン ODIN を用いて評価実験をおこない，実際にランキングの質が向上することを示す．

A Searching and Ranking Scheme using Hyperlinks and Anchor Texts

Kazuhiro KAZAMA, Masanori HARADA, Shin-ya SATO
Nippon Telegraph and Telephone Corporation
Network Innovation Laboratories

This paper describes a searching and ranking scheme using hyperlinks and their associated anchor texts in order to improve a search engine's ranking quality. The results of our experiments with this scheme show an improvement of ranking quality for a search engine called "ODIN".

1　はじめに

World Wide Web の情報空間の内容の充実と拡大に伴って，目的の情報を探す手がかりとしてのサーチエンジンの重要性が高まっている．しかし，情報検索に対する充分な知識を持たないサーチエンジンのユーザが，膨大で多様な情報の中から，目的とする情報を探し出すことは容易ではない．

現在，ハイパーリンク構造を解析して求めた Web ページの重要度を利用する手法が提案されているが，検索質問によっては必ずしも適切なランキングをおこなえない問題がある．

本稿では，リンク元ページのアンカーテキストをリンク先ページのテキストの一部として索引付けして検索することで，検索質問に適合したハイパーリンクだけを考慮できるようにすると共に，サーバ内アンカーテキストとサーバ外アンカーテキストの性質を考慮して異なる重みを与えることで，適切な検索結果のランキングを実現する手法について述べる．さらに，サーチエンジン ODIN を用いて実験をおこない，その有効性を示す．

2　サーチエンジンとハイパーリンク

2.1　サーチエンジンの問題

Web 空間の情報を探すためのサーチエンジンを一般の情報検索システムと比較した場合に，いくつかの問題が存在する．

一番目は，Web 文書は，新聞記事データベー

スなどと比べると，用語や文体が統一されていなかったり．ハイパーテキストとしてのさまざまな文書構造が存在するために，情報の取り扱いが難しいことである．このために，膨大な検索結果中にさまざまの品質の情報が順不同で混在することになることが多いが，実際には検索結果の上位しか見れないので，目的の情報を探し出せるとは限らない．

二番目は，サーチエンジンのユーザは，サーチャのような情報検索についての知識を持つ人が非常に少ないために，論理演算子やオプション機能をほとんど使わないだけでなく，ごくわずかの検索語しか使用しない傾向にあることである．たとえば，Silverstein らは，AltaVista のログを解析し，ユーザは 1 ～ 3 語 (平均 2.35 語) を検索に使用していると述べている [1]．複合語が空白で分割されない日本語では検索語数はさらに少なくなり，ODIN のログを解析した結果では，平均検索語数が 1.42 語であり，1 ～ 2 語が 91.4% を占めている [2]．

つまり，サーチエンジンのユーザは，Web 空間の膨大で多様な情報に対して，わずか 1 ～ 2 語を入力するだけで検索しようとしているわけである．

この厳しい要求に応えるためには，既存の IR (Information Retrieval) の技術のように検索質問と文書の適合性を考えるだけでは不充分である．

2.2 ハイパーリンクの利用

文書 A が文書 B にリンクした場合には，文書 A の著者は文書 B に何らかの価値があると考えていると推測できる．この仮定に基づいてハイパーリンク構造を解析して，Web 空間における文書の重要度を求める手法として，現在 Link Popularity，HITS，PageRank，Cocitation アルゴリズムなどが提案されており，検索結果のランキングに利用されている．

Link Popularity **Link Popularity** は，数多くリンクされている Web ページほど重要だと見なす手法である．たとえば，検索された Web ページへの被リンク数や，検索された Web ページが存在するサーバへの総被リンク数の多い検索結果のスコアを高くする手法が使われている．

HITS Kleinberg は，ある特定のトピックに関する情報源であるオーソリティ (authority) と，オーソリティへのリンクの集合であるハブ (hub) という 2 種類の Web ページに対して，よいオーソリティは多くのよいハブからリンクされ，良いハブは多くの重要なオーソリティをリンクするという相互依存する関係を求めることで，検索結果の質を改善する **HITS** (Hypertext Induced Topic Search) を提案した [3]．

PageRank Page らは，多くの良質な Web ページからリンクされている Web ページは，良質な情報源であると考え，Web のリンクをランダムに辿る “random surfer” 行動モデルに基づいて，Web ページを閲覧する確率を計算することで，Web ページの重要度を示す **PageRank** を求める手法を提案した [4, 5]．

Cocitation アルゴリズム Dean らは，内容の類似している Web ページをリンク解析で発見する手法として，2 つの Web ページを同時に引用している Web ページ群を求めて，その数により内容が類似した Web ページを見つけ出す **Cocitation アルゴリズム**を提案した [6]．

2.3 トピックドリフト問題

これらのハイパーリンクを利用する手法では，検索結果の上位に検索語と関連のない Web ページがランキングされることがある．これをトピックドリフト問題 (topic drift problem) と呼ぶ [7]．

この原因は，それぞれのハイパーリンクがどのような意図に基づいているかを考慮しないために，一般的な著名サイトがランキング上位になりやすいだけでなく，プログラムが機械的に生成したハイパーリンクの影響を受けやすい．

図 1: A 要素のアンカー例

```
For more information about W3C, please consult the
<A href="http://www.w3.org/">W3C Web site</A>.
```

さらに，これらの手法は検索質問とは無関係なので，一般的なトピックを表す検索語 A と専門的なトピックを示す検索語 B を用いて，“A and B” のように AND 検索をおこなった場合には，A のトピックに偏りやすい．

検索語と関連があるハイパーリンクを選択する手法として，同じトピックへのハイパーリンクは Web ページ内で近隣に配置される傾向があることを利用して HITS を改良した Companion アルゴリズムが提案されている [6]．これは例外も多く根本的な解決ではないだけでなく，検索質問を考慮しないために複数の検索語が指定された場合は適切な結果が得られるとは限らない．

また，検索質問と Web ページの内容の適合性を求めて，不適切なノードを除外したり，影響を制御することで HITS を改良する手法も提案されている [7]．これは計算コストがかかるだけでなく，検索語は含まれないが内容は適切である Web ページが除外される危険がある．

3 アンカーテキストの検索とランキング

3.1 アンカーテキスト

ハイパーリンクはアンカー (anchor) と呼ぶ 2 つの端を持ち，ソース・アンカー (source anchor) から任意の Web リソースに対するデスティネーション・アンカー (destination anchor) へとリンクされる [8]．たとえば，図 1に示す A 要素の例では，ソース・アンカーが，“W3C Web site” という文字列であり，デスティネーション・アンカーが，http://www.w3.org/である．

特に，ソースアンカー中のテキストをアンカーテキスト (anchor text) と呼ぶ．なお，本稿では，ソース・アンカー中の IMG 要素の ALT 属性，イメージマップ中の AREA 要素の ALT 属性もアンカーテキストとして扱うことにする．

3.2 アンカーテキストとサーチエンジン

アンカーテキストは，リンク先の内容を的確に表現したり，リンクした意図を反映することが多いという特徴があると考えられるので，World Wide Web Worm (WWW) や Google などのサーチエンジンで古くから用いられてきた [9, 4, 5]．

鷲崎らによるアンカーテキストの解析結果によると，情報検索で重要な簡潔な名詞，複合名詞，名詞句が主に使われていることがわかる [10]．

さらに，アンカーテキストを使うことで，テキストを含まない画像などの Web リソースや，Web ロボットが収集していない Web ページも検索できる．

3.3 アンカーテキストの分類

アンカーテキストは，Web 文書の作者自身が記述した場合と，作者以外の人間が記述した場合に分類でき，それぞれ異なる性質を持っていると考えられる．

前者は，Web 文書を構成するほとんどのページに対して付加されるが，作者自身が記述しているので表記が統一されている．サイトの構成によっては Web 文書が相互にリンクしていたり，「次」などの内容とは関連のない単語をツールが自動生成していることも多い．

後者は，Web 文書が重要または関連があると判断された場合に，目次となるような代表的なページに対して付加される．ただし，他人が記述するために，略語や俗称などの多種多様な表記が使われており，変換ミスなどの間違いもある．

なお，実際には META 要素の author プロパティや LINK 要素の made プロパティを使っても Web 文書の作者を正しく判定することが難しいので，本稿では，それぞれ同一サーバ内に存在する場合と，異なるサーバに存在する場合で近似することにする．以後，前者をサーバ内アンカーテキ

スト，後者をサーバ外アンカーテキストと表記する．

3.4 アンカーテキストの検索とランキング

本稿では，リンク元ページのアンカーテキストをサーバ内アンカーテキストとサーバ外アンカーテキストに分けてから，それぞれリンク先ページのテキストの一部として索引付けし，検索時にそれぞれの性質を考慮した重みを用いてスコアを求めることで，ランキングの質を向上する手法を提案する．

ODIN に用いている全文検索エンジン Jerky では，同一索引中に文書の情報を特性に合わせて別々に格納し，それらを組み合わせて検索することができる．

そこで，ロボットで収集した HTML ファイルから，本文，タイトル，META 要素の keywords プロパティ，META 要素の description プロパティ，そして，その HTML ファイルに対するサーバ外アンカーテキストの集合とサーバ内アンカーテキストの集合の 6 種類のプロパティを抽出し，索引を作成する．ここで，それぞれのプロパティの集合 P を，次のように定義する．

$$P = \{text, title, keywords, description, eanchor, ianchor\}$$

検索時には，まず入力 s を解析し，検索語のリスト $k_0, k_1, ..., k_n$ を求める．

$p \in P$ に対して，$tf_{k_i}^p$ を，プロパティ p 中に検索語 k_i が出現する頻度とする．たとえば，$tf_{k_i}^{eanchor}$ は，サーバ外アンカーテキスト集合中の検索語 k_i の出現頻度である．

一方，各プロパティに対して，重み w^p を決定し，検索語 k_i の TF (Term Frequency) 値 TF_{k_i} を，以下のように定義する．

$$TF_{k_i} = \sum_{p \in P} w^p tf_{k_i}^p$$

検索結果スコアは，このように求められた TF_{k_i} の値に基づいて，TF・IDF 法に基づいて計算する．

これよりわかるように，TF_{k_i} の値は，重みの組 $\mathbf{w} = w^p$ に応じて変化することになる．

3.5 重みの設定

外部アンカーテキストの出現頻度は，インターネットのユーザにとっての重要性を示す．従来のハイパーリンクを利用した手法が，Web ページの人気を求めるものであるが，外部アンカーテキストの出現頻度を求めることは，検索質問に対して適切だと思われる Web ページを推薦した結果を集計し，一般的な合意を求めることと考えることができる．

内部アンカーテキストの出現頻度は，その Web ページが属する Web 文書にとっての重要性を示す．これは文書構造によって変化するために，必ずしも数が多ければ良いとは限らない．

本稿では，従来の検索質問と文書の適合性に加えて，この 2 つの重要性を考慮してスコアを計算することになる．

ただし，一般的な合意を表し，信頼性の高い外部アンカーテキストの出現頻度が一番重要だと考えられるので，$w^{eanchor}$ を w^{text} や $w^{ianchor}$ より大きく設定している．

3.6 トピックドリフト問題の解消

本手法では，検索語を含むアンカーテキストを持つハイパーリンクだけが反映されるので，“Yahoo! Japan” などの極度に被リンク数が多いサイトへのハイパーリンクや，プログラムが機械的に生成した可能性が高い「Next」のようなアンカーテキストが付加されたハイパーリンクに影響されない．

さらに，従来のスコア計算と，ハイパーリンクの被リンク数に基づくスコア計算を同一レベルで統合できたので，本手法ではトピックドリフト問題が発生しない．

表 1: アンカーテキストの出現頻度

外部アンカーテキスト	出現頻度
NTT	187
NTT Home Page	22
NTT ホームページ	21
日本電信電話株式会社	20
http://www.ntt.co.jp/	17
日本電信電話 (株)	16
日本電信電話	13
日本電信電話 (NTT)	8
www.ntt.co.jp	6
NTT Homepage	6
Nippon Telegraph and Telephone Corporation	5
NTT のホームページ	5
NIPPON TELEGRAPH AND TELEPHONE CORPORATION	5
内部アンカーテキスト	**出現頻度**
HOME	26
http://www.ntt.co.jp/	2
NTT(持ち株会社)	1
http://www.ntt.co.jp	1

4 評価

4.1 アンカーテキストの解析

サーバ内・サーバ外アンカーテキストの性質に対して，www.ntt.co.jp を取り上げて，JP ドメインから収集した 810 万 9330 ページの HTML ファイルを解析した結果について述べる．

www.ntt.co.jp からは 2,710 ページを収集した．未収集の Web ページを含めると 2,744 ページに対するハイパーリンクが存在する．サーバ内では 2,527 ページがリンクされているが，サーバ外からは全体の 10.5% の 289 ページしかリンクされていない．後者を調べると，情報のクラスタ中の目次となる Web ページに集中してリンクしていることがわかる．

アンカーテキスト数は，サーバ内が 7,080 個，サーバ外からが 3,123 個，計 10,203 個である．ただし，同じ文字列の重複が多く，単なる半角・全角の違いやスペース文字の違いもあるので，正規化した後は 5,349 種類となる．そのうち，サーバ内が 1,310 種類，サーバ外からが 4,052 種類になる．

さらにトップページだけに注目すると，488 個，112 種類となる．そのうち，サーバ内アンカーテキストが 30 個，4 種類，サーバ外アンカーテキストが 457 個，109 種類となり，これを頻度の多い順に表 1に示す．

この結果から，サーバ内アンカーテキストは比較的同じ文字列が使われているが，サーバ外アンカーテキストは多種多様な文字列が使われていることがわかる．

みかか (「み」と “N”，「か」と “T” が同じキーに割り当てられていることから，NTT を示す) や証券番号，URL など，NTT という会社を指すために多彩な表現が使われていたり，「持株会社」や「NTT グループ」などリンクした目的が異なる場合が存在することがわかる．

これらの結果から，サーバ外アンカーテキストは，目次となる Web ページの順位を向上させるだけでなく，表記の揺らぎや視点の違いを吸収できることがわかる．

4.2 アンカーテキストの検索数

本手法では，検索語と一致したアンカーテキストを持つハイパーリンクだけが反映されるので，ハイパーリンク数がどの程度確保できるかどうかが重要である．

そこで，2000 年 1 月 1 日から 6 月 30 日までの間に ODIN で使用された検索語を使用頻度に応じて順位付けした後に，50 間隔で抽出した 100 個の単語の全検索結果数と，外部アンカーテキストと内部アンカーテキストで検索された結果数を求め，全検索結果数順に並べた結果を図 2に示す．

内部アンカーテキストは全検索結果の 4.6% 程度，外部アンカーテキストは 0.5% 程度である．外部アンカーテキストの検索結果数は，該当するアンカーテキスト数よりかなり低いが，これは特定の Web ページに集中しているからである．

また，適切と思われるハイパーリンクが存在しても，外部アンカーテキストの検索結果が常に得られるとは限らないが，このような場合にアンカーテキストを補完する研究として，Chakrabarti らの anchor window を用いてアンカー周辺のテキストを含める研究や，鷲崎らの経験則を用いてアンカーテキストの範囲を拡張する研究が存在す

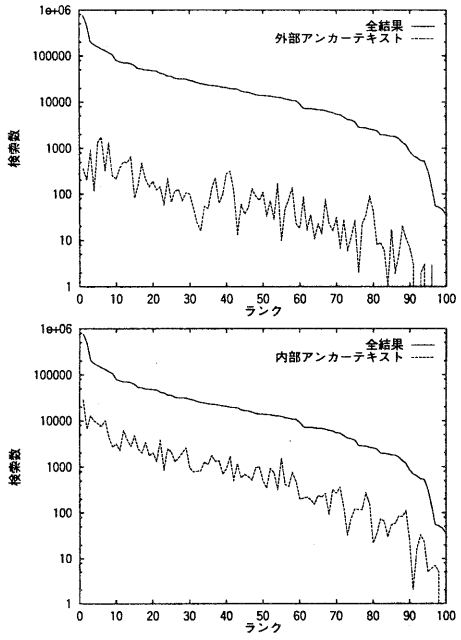

図 2: アンカーテキストの検索数

る [11, 10].

しかし，図2からわかるように，このような場合は検索結果数自体が少なく，検索結果を閲覧するのはそれほど困難ではない．さらに，そのような場合には，外部アンカーテキストより数の多い内部アンカーテキストの検索結果を用いて，ランキングを改善することができる．

4.3 オフィシャルサイトの順位

本手法のランキングの妥当性を示すために，検索語とサイトの関係を一意に決定できるオフィシャルサイトの検索結果中の順位を調べる．

まず，Yahoo Japan! の各カテゴリから，JP ドメイン内にオフィシャルサイトを持っている登録名とその URL の組を 1 カテゴリにつき 5 つ抜き出して，登録名で検索した場合のオフィシャルサイトのトップページの順位と，オフィシャルサイト中の Web ページの一番よい順位を求めた．オフィシャルサイトのトップページは，URL の最後が "/" か，またファイル名の先頭が "index" で最後が ".html" か ".htm" かどうかで識別した．

この実験では，$w1$ (= [1, 10, 5, 2, 0, 0])，$w2$ (= [1, 10, 5, 2, 0, 8])，$w3$ (= [1, 10, 5, 2, 8, 0])，$w4$ (= [1, 10, 5, 2, 8, 8])，$w5$ (= [1, 10, 5, 2, 12, 1]) という 5 つの重み条件を用いた．

この実験結果を，表2に示す．この結果から，サーバ内アンカーテキストを使用した $w2$ とサーバ外アンカーテキストを使用した $w3$ とも順位が向上するが，$w3$ の効果の方が著しい．$w3$ では，トピックドリフト問題が顕著に出やすい個人的なサイトや，No.4 のようにトップページに画像を多用し検索語が含まれない場合も適切にランキングできている．

ただし，WWWW や Google のようにサーバ外とサーバ内アンカーテキストを区別しない $w3$ は，$w2$ より実験結果が悪くなるが，$w5$ のようにサーバ内の重みを減らし，サーバ外の重みを増やすことで改善でき，ほぼ確実に 1 ページ目 (1 ～ 10 件) に表示できるようになる．

順位が 1 位にならない原因は，次のように分類できる．

1. オフィシャルサイト内の他の Web ページが上位にリストアップされる．No.48 ではカルピス・ウォータ，No.50 ではヤクルトスワローズが上位に来る．

2. オフィシャルサイトでないが，妥当なホームページが上位にリストアップされる．No.24 では同一組織の他サイトで，No.44 はサイトの場合も多い．

3. トップページが複数存在する．No.6 では "mthome_htm" が，No.36 では "jhome.html" が上位に来るが，トップページと内容は同じである．

オフィシャルサイトを検索結果の先頭やランキング上位に表示するために，他のデータベースを併用するサーチエンジンもあるが，本手法は Web ロボットで収集したデータだけで任意のトピックに対して同じ効果が得られる．

また，高野らのPageRankの研究では必ずしもよい結果が得られていないにもかかわらず，Googleでよい結果が得られるのはアンカーテキストを併用しているからだと考えられるが，本手法はアンカーテキストだけで同等の結果が得られる[12, 4].

4.4 今後の課題

評価実験の拡大 オフィシャルサイトの検索は，サーチエンジンの使用方法の一部でしかない．複雑な検索質問や，最適解が一意に決まらない場合に対しても，評価する必要がある．

重みの決定 上記の理由から，今回示した重みは最適値とは限らない．そこで，重みについての理論的分析や，自律学習による最適化を検討している．

情報更新への対応 新しいバージョンのソフトウェアのWebページが新設されたり，URL変更時に古いWebページが削除されずに移転情報を示すために残されている場合には，長い間古い方を上位に表示されることがある．

表2のNo.44も翌月には1位になった事実からわかるように，リンクするのは簡単で被リンク数の増加も比較的早い．しかし，常に実体と合うように管理されるわけではなく，大幅な遅延が生じるので，これに対応する必要がある．

5 おわりに

本稿では，ハイパーリンクだけでなく，それに関連づけられたアンカーテキストを使って検索し，さらにサーバ内アンカーテキストとサーバ外アンカーテキストを区別することで，トピックドリフト問題を回避し，適切なランキングを実現する手法について述べた．

さらに，実際にODINの運用データを用いて評価実験をおこない，その有効性を証明した．

参考文献

[1] Silverstein, C., Henzinger, M., Marais, J., Moricz, M.: Analysis of a very large AltaVista query log, Technical Report 1998-014, Compaq Systems Research Center, 1998.

[2] ODIN: http://odin.ingrid.org/

[3] Kleinberg, J. M.: Authoritative sources in a hyperlinked environment, the Journal of the ACM, 1999.

[4] Brin, S., Page, L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine, The 7th International World Wide Web Conference, 1998.

[5] Google: http://www.google.com/

[6] Dean, J., Henzinger M. R.: Finding Related Pages in the World Wide Web, The 8th International World Wide Web Conference, 1999.

[7] Bharat, K., Henzinger, M. R.: Improved Algorithms for Topic Distillation in Hyperlinked Environments, Proceedings of the 21st International ACM SIGIR Conference on Research and Development in Information Retrieval, 1998.

[8] Raggett, D., Hors, A. L., Jacobs, I. : HTML 4.01 Specification, 1999.

[9] McBryan, O. A.: GENVL and WWWW: Tools for Taming the Web, The 1st International Conference on the World Wide Web, 1994.

[10] 鷲崎誠司, 村本達也: ハイパーリンクの構造を利用した検索結果の選択手法, SIGFI 55-10, 1999.

[11] Chakrabarti, S., et. al. Automatic resource list compilation by analyzing hyperlink structure and associated text, Proceedings 7th International World Wide Web Conference, 1998.

[12] 高野元, 久保進也: サイテーション・エンジン: リンク解析を用いたWWW検索ランキングシステム, データベースシステム 120-2, pp.9 –16, 2000.

表 2: 重みと順位

	検索語	$w1$		$w2$		$w3$		$w4$		$w5$	
1	富士重工業	349	299	46	46	1	1	1	1	1	1
2	本田技研工業	32	16	52	31	1	1	1	1	1	1
3	日産自動車	39	23	97	85	1	1	1	1	1	1
4	ポルシェジャパン	58	41	58	41	1	1	1	1	1	1
5	フィアットオートジャパン	23	23	23	23	1	1	1	1	1	1
6	運輸省	32	29	102	97	2	1	2	1	2	1
7	厚生省	1218	1193	207	183	1	1	1	1	1	1
8	首相官邸	22	20	1	1	1	1	1	1	1	1
9	通商産業省	9002	8573	9037	8608	1	1	1	1	1	1
10	郵政省	124	122	1	1	1	1	1	1	1	1
11	産經新聞	3946	3887	7	6	1	1	7	6	1	1
12	中日新聞	919	878	1	1	2	2	1	1	1	1
13	日本経済新聞	16630	16546	135	121	1	1	1	1	1	1
14	毎日新聞	622	622	762	762	1	1	1	1	1	1
15	読売新聞	18256	16309	18263	16345	1	1	1	1	1	1
16	安達祐実	679	679	5	5	2	2	2	2	1	1
17	榎本加奈子	1315	1315	1369	1369	1	1	1	1	1	1
18	小泉今日子	43	41	76	72	2	2	2	2	2	2
19	西城秀樹	855	855	859	859	1	1	1	1	1	1
20	松雪泰子	670	670	2	2	1	1	1	1	1	1
21	東京大学	1631	19	620	488	1	1	1	1	1	1
22	一橋大学	70	63	15	15	1	1	1	1	1	1
23	慶應義塾大学	43	43	106	106	1	1	1	1	1	1
24	青山学院大学	4145	3343	4157	3359	5	5	14	14	6	6
25	國學院大学	472	320	1	1	1	1	1	1	1	1
26	東京国立近代美術館	99	98	2	2	1	1	1	1	1	1
27	サントリー美術館	2	2	3	3	1	1	1	1	1	1
28	秋田県立近代美術館	13	13	13	13	2	2	4	4	2	2
29	練馬区立美術館	3	3	4	4	4	4	4	4	5	5
30	徳川美術館	2	1	1	1	1	1	1	1	1	1
31	ZDNet Japan	50	32	56	34	1	1	1	1	1	1
32	HotWired Japan	20	19	20	19	1	1	1	1	1	1
33	インターネット ASCII	151	151	3	3	1	1	1	1	1	1
34	日経コンピュータ	5	5	10	10	1	1	1	1	1	1
35	Delphi マガジン	6	5	20	19	1	1	1	1	1	1
36	理化学研究所	40	38	8	5	2	1	2	1	2	1
37	科学技術振興事業団	281	279	47	46	1	1	1	1	1	1
38	鹿児島県工業技術センター	1	1	1	1	1	1	1	1	1	1
39	日本コンピュータ支援外科学会	2	2	7	6	1	1	1	1	1	1
40	基盤技術研究促進センター	3	2	7	5	1	1	1	1	1	1
41	ラサール石井	4	3	8	7	1	1	2	1	1	1
42	西村雅彦	674	674	674	674	2	2	3	3	1	1
43	藤田弓子	2	2	2	2	1	1	1	1	1	1
44	松たか子	149	129	180	160	12	12	26	26	10	10
45	杉田かおる	17	16	3	3	1	1	1	1	1	1
46	DyDo	230	220	1	1	1	1	1	1	1	1
47	POM ジュース	32	1	1	1	1	1	1	1	1	1
48	カルピス	61	54	7	5	4	3	4	2	3	2
49	コカ・コーラ	42	41	2	2	1	1	1	1	1	1
50	ヤクルト	290	289	38	32	2	1	2	1	2	1
	平均順位	1267.48	1160.18	742.4	673.7	1.58	1.5	2.22	2.08	1.5	1.42