

学士學位論文

PageRank を利用した blog 人気度の判定

茨城大学工学部

システム工学科

執筆者：茂木啓悟

指導教官：新納浩幸

平成 17 年 3 月 2 日

# 目次

第1章	はじめに	5
1.1	研究概要	5
1.2	本論文の構成	6
第2章	blog とその基盤技術	7
2.1	blog の概要	7
2.1.1	コンテンツ更新処理の自動化	8
2.1.2	RSS データの発信	8
2.1.3	読者参加型 (comments)	8
2.1.4	別の blog との相互参照 (trackback)	9
2.2	RSS フォーマット	10
2.2.1	RSS0.9	10
2.2.2	RSS0.91	10
2.2.3	RSS1.0	10
2.2.4	RSS0.92 / RSS 2.0	11
2.2.5	RDF(Resource Description Framework)	11
2.3	XML	14
2.3.1	XML(eXtensible Markup Language)	14
2.3.2	SGML(Standard Generalized Mark-up Language)	16
2.3.3	XSL(eXtensible Stylesheet Language)	16
2.3.4	XSLT(XSL Transformations)	17
2.3.5	XLink/XPointer(XML Linking Language/XML Pointer Language)	17
第3章	blog のランキング手法	18
3.1	単一プロバイダ内の blog ランキング手法	18
3.1.1	livedoorBlog の blog ランキング方法 ( <a href="http://blog.livedoor.com/">http://blog.livedoor.com/</a> )	18

3.1.2	gooBLOG の blog ランキング方法 ( <a href="http://blog.goo.ne.jp/">http://blog.goo.ne.jp/</a> ) . . .	20
3.1.3	AmebaBlog の blog ランキング方法 ( <a href="http://ameblo.jp/">http://ameblo.jp/</a> ) . . . . .	21
3.2	プロバイダに依存しない blog ランキング手法 . . . . .	22
3.2.1	Blog Ranking( <a href="http://blog.with2.net/">http://blog.with2.net/</a> ) . . . . .	22
3.2.2	週刊ブログ王 ( <a href="http://www.doramix.com/rank/">http://www.doramix.com/rank/</a> ) . . . . .	22
<b>第 4 章</b>	<b>Web ページ重要度を測る手法</b>	<b>23</b>
4.1	PageRank . . . . .	23
4.1.1	基本的な計算式 . . . . .	24
4.1.2	行列計算による方法 . . . . .	24
4.1.3	リンクを用いないアクセスの考慮 . . . . .	25
4.2	HITS(Hypertext Induced Topic Search) . . . . .	27
4.2.1	HITS アルゴリズム . . . . .	27
4.2.2	オーソリティ . . . . .	27
4.2.3	ハブ . . . . .	27
<b>第 5 章</b>	<b>PageRank を利用した blog 評価方法</b>	<b>29</b>
<b>第 6 章</b>	<b>実験</b>	<b>31</b>
<b>第 7 章</b>	<b>考察</b>	<b>39</b>
<b>第 8 章</b>	<b>おわりに</b>	<b>40</b>
<b>第 9 章</b>	<b>謝辞</b>	<b>41</b>
<b>付録 A</b>	<b>プログラムソースリスト</b>	<b>43</b>

## 目次

2.1	Blog トップページ例 ( <a href="http://alejandro-rivero.ameblo.jp/">http://alejandro-rivero.ameblo.jp/</a> ) . . . . .	12
2.2	blog の XML 表示 . . . . .	13
3.1	livedoorBlog トップページ . . . . .	19
3.2	gooBlog トップページ . . . . .	20
3.3	AmebaBlog トップページ . . . . .	21
4.1	PageRank の概念図 . . . . .	24
4.2	リンク例 . . . . .	25
4.3	ハブとオーソリティの概念図 . . . . .	28
5.1	google による被リンク数の計測 . . . . .	30
6.1	livedoorBlog ランキングの実験結果のグラフ . . . . .	38

## 表目次

2.1 XML の例 . . . . .	16
6.1 livedoorBLOG - 1 ( <a href="http://www.livedoor.com/">http://www.livedoor.com/</a> ) . . . . .	32
6.2 livedoorBLOG - 2 ( <a href="http://www.livedoor.com/">http://www.livedoor.com/</a> ) . . . . .	33
6.3 gooBLOG( <a href="http://blog.goo.ne.jp/">http://blog.goo.ne.jp/</a> ) . . . . .	34
6.4 アメーバ BLOG( <a href="http://ameblo.jp/">http://ameblo.jp/</a> ) . . . . .	35
6.5 人気 blog ランキング ( <a href="http://blog.with2.net/">http://blog.with2.net/</a> ) . . . . .	36
6.6 週刊 blog 王 ( <a href="http://www.doramix.com/rank/">http://www.doramix.com/rank/</a> ) . . . . .	37

# 第1章 はじめに

## 1.1 研究概要

blogとは、正確にはWeblogと言われるもので、サイト運営者個人の専門分野に関する知識や興味関心のある情報が集まったWeb上に残る記録として定義される。また、コンテンツを作成する特性も含めてblogと呼ぶ場合もある。blogは、個人の日記を記すことよりも、話題となっているインターネット上のニュースを引用し、コメントや批判を加えたり、新しい視点を提供するといった働きを持っている。それゆえ、ニュースサイトへのリンクを簡単に貼ることができるので、あまり目に止めなかったニュースに世間の注意を引き付ける役割りを果たすことができ、インターネットを用いたニュース報道の新しい形として注目されている[1]。

流行を知る上で参考となるのがランキングであるが、既存のblog人気ランキングはプロバイダごとに登録した各blogサイトで、主にアクセス数によって決められている。そこで本研究では、いろいろなプロバイダ間のblogにおける統一ランキングを作る上で必要となるblogの人気評価手法を提案する。

統一ランキングを作るために、まず各プロバイダのblogについて評価を行い、各プロバイダの持つランキングに近づけることを目標とする。各プロバイダはそれぞれアクセス数などそれぞれ独自のblog評価基準を持っており、それらは公表されていない。そこで本実験では人気度を測る基準として第三者でも測ることができる被リンク数を用いる。そして、Webの重要度を測る手法としてgoogleで用いられているPageRankアルゴリズムを用い、各blogに対するリンク数およびリンクしているページの重要度をblog人気度の評価基準とし、各blogに対して評価を行う。

## 1.2 本論文の構成

本論文では始めに、世間に対する blog の位置付け、blog の特徴、blog を構成する技術についての説明を行う。次に、プロバイダ内の blog について人気度を評価する手法、blog ランキングに登録した blog に対して人気度を評価する手法の2つの手法について述べる。その後に Web ページの重要度を測るアルゴリズムの例として PageRank アルゴリズム、HITS アルゴリズムについての説明を行い、blog 人気度を測る手法として PageRank アルゴリズムを利用した blog 評価を提案する。PageRank を利用した blog 人気度の評価の実験の解説と結果を示し、その結果に対する考察を行う。

## 第2章 blogとその基盤技術

### 2.1 blogの概要

日記形式のWebサイトのことをWeblog（ウェブログ）、または単にblog（ブログ）と呼んでいる。インターネットユーザーたちがインターネット上で見つけたおもしろいコンテンツやニュースに対してメモ（ログ）をつけだしたのがblogの始まりだと言われ、これらの記録にサイト運営者のコメントがつけられ運営者を中心としたユーザー間でコミュニケーションを行うための便利なコンテンツとなった。

blogを公開するにはWebサーバを用意し、そのWebサーバ上にブログシステムを導入する必要があるが、この作業は技術に詳しい人でなければ難しく、回線やサーバなどに費用もかかる。このため、ポータルサイトやインターネットサービスプロバイダ（ISP）は自社ユーザへのサービスの一環として、自社でblogシステムを運用し、ユーザが簡単にblogを公開できるサービスを始めた。多くのサービスは無料で提供されており、Webブラウザさえあれば簡単に自分のblogを開設することができる。blogサービスの事業者は、追加機能を有料で提供したり、ページに広告を入れることで収益を得ている。ランキングを公表したり、アクセスの多いblogのオーナーには賞金を出すことで集客に努めている。

最近では携帯電話を使って、短い文章やカメラ機能で撮影した写真などを頻繁に投稿するという形態の「moblog」というものが増えてきている。moblogは「モバイル（mobile）と「ブログ」（blog）を合わせた造語で、1日の決まった時間にパソコンを使ってまとまった量の文章を投稿することの多い通常のblogに対し、独特のライブ感覚やテンポの良さがあり、内容的にも生活に密着した身近な話題が取り上げられている例が多い。出かけ先で簡単に投稿できる点が最大の魅力である。

多くのblogには読者から容易にフィードバックできる機構が用意されており、話題ごとにコミュニティが形成されている。また、別のblogの気になる記事へのリンクを張る相互参照も盛んで、コミュニティ内で話題のトピックには自然と注目が集まるメカニズムができています。最近では、テレビなどでは紹介されないロー

カル情報など様々な情報がblogを通じてやりとりされ、マスメディアが後追いでそのトピックを取り上げるという現象も起こっている。そういった面からもblogは新しいメディアとして注目され、世間一般への影響力を持つようになった。

blogは次のような特徴を備えている。

### 2.1.1 コンテンツ更新処理の自動化

blogのコンテンツを更新する際は、ツールやサービスの専用フォームに入力して定型レイアウトに流し込むことでHTMLを自動編集する方法が一般的である。Webブラウザなどからアクセスして記事を投稿すると適切なHTMLファイルが自動的に生成される。

### 2.1.2 RSSデータの発信

RSS(Rich Site Summary)とは、Webサイトの見出しや要約などのメタデータを構造化して記述するXMLベースのフォーマットで、主にサイトの更新情報を公開するために使われている。RSSで記述された文書には、Webサイトの各ページのタイトル、アドレス、見出し、要約、更新時刻などの情報を記述することができる。RSS文書を用いることで、多数のWebサイトの更新情報を統一的な方法で効率的に把握することができる。指定したサイトのRSS情報を取り込んで更新状況をまとめたWebページを生成するアンテナ(巡回)ソフトや、デスクトップに指定したサイトの更新情報を表示するティッカーソフトなどが開発されている。blogではRSSに記述されたWebサイトの要約データを出力する。

### 2.1.3 読者参加型(comments)

blogでは著者から記事が一方向的に投稿されるのではなく、訪問者がblogの各記事で掲示板のようにコメント(comments)を付けることができる。記事単体のページにあるフォームから、名前やコメントを入力し書き込むことができる。このコメントの機能により、著者と読者がコミュニケーションをはかったり、特定のトピックスについて電子掲示板で多人数で論議することが容易となる。

#### 2.1.4 別の blog との相互参照 (trackback)

ある blog を訪れた人が自分の blog 上で相手の blog へ trackback (トラックバック) という処理を行うと、リンク先の相手にリンクを張ったと通知される。つまり、リンク元への逆リンクが自動的に行われる。

blog 作者が別の blog の記事を参照して自身のサイトにコメントを掲載するような場合、元の記事へのリンクを張るのが一般的だが、単にリンクしただけでは元の記事の作者はどこからどうリンクされているのかを容易に知ることはできない。トラックバックはリンク元サイトに「このような記事からリンクを張った」という情報を通知する仕組みで、リンク元記事の URL やタイトル、内容の要約などが送信される。トラックバックされたサイトではこの情報を元に「この記事参照している記事一覧」を自動的に生成することができる。

相手に送信されるトラックバック通知のことを「トラックバック・ピング」(trackback ping) と呼び、通知の送信先を「トラックバック URL」(trackback URL) という。多くの blog の記事には隅に「この記事へのトラックバック URL」が記載されている。また、トラックバックはコメントとは異なり相互参照性が高い。

## 2.2 RSS フォーマット

### 2.2.1 RSS0.9

最初の RSS である RSS 0.9 は、RDF Site Summary として、1999 年に米国 Netscape Communications 社が自社のポータルサイト「My Netscape」において、「チャンネル」の詳細を記すために策定したものである。RDF 構文を用いたことから、RDF Site Summary と呼ばれる。

間もなく Netscape は RSS から手を引きその仕様を UserLand Software に譲渡した。その後、要素を拡張した独自の XML フォーマットで記述される RSS 0.91 が策定された。

### 2.2.2 RSS0.91

Rich Site Summary と改名された RSS 0.91 は、RSS 0.9 に要素を拡張する目的で作られた。RDF を用いず、独自の XML で記述されている。

UserLand Software の ScriptingNews から著作権、日付情報などいくつかの要素を取り入れ拡張された。それまでの RSS 0.9 より多くの情報を配信できるようになったため、Rich Site Summary と呼ばれ、その後派生した RSS 0.92, RSS 2.0 のベースとなった。

RDF 0.91 の登場以降、RSS が持つ「コンテンツ配信」機能に対しての需要がさらに高まった。そのためよりリッチなコンテンツ配信を目指そうとする制作者が、独自の要素を RSS に追加してしまうなど、フォーマットの拡張における混乱が生じた。

### 2.2.3 RSS1.0

RSS 1.0 は XML 名前空間を利用した拡張性を重要な特徴としている。名前空間を宣言して Creative Commons や FOAF といったポピュラーな語彙、あるいは独自の要素型を追加することもできる。より高い相互運用性のために、RSS-DEV Working Group で拡張語彙の標準「モジュール」を定めている。

RSS 1.0 は 0.9 時代に使われた RDF を再び採用し、RSS が持つ「メタデータ記述」としての側面を主眼に置いたフォーマットとなっている。

RSS コアモジュールの他に公式なモジュールとして、Dublin Core モジュール、Syndication モジュール及び Content モジュールが定められた。これにより RSS 0.9 で不満であった語彙の乏しさを解消させ、さらにコンテンツ配信手段として RSS 1.0 を採用する道を残すものとなった。

RSS 1.0 の登場により、メタデータ記述技術としての RSS の中核語彙及び拡張性が保証されるようになった。しかし RDF を再び採用したこと、モジュールによる XML 名前空間の複雑化は RSS 配信者の不満の種となり、RSS 0.91 系のフォーマットを拡張する動きが再びみられることとなった。

日本においては、普及している RSS のほとんどが RSS 1.0 を用いているため、RSS は RDF Site Summary として認識されている。

#### 2.2.4 RSS0.92 / RSS 2.0

拡張をオプションとして提供し、かつ RSS 0.91 への互換性を持たせる方法が提案され、それを受けて 2000 年 12 月に UserLand Software から RSS 0.92 が発表された。

UserLand はその後も互換性を維持したまま RSS 0.93, RSS 0.94 という拡張を続けたが、2002 年 8 月に RSS 0.91 から RSS 0.94 までのすべてのフォーマットに対する互換性を保証した RSS 2.0 を策定し、これを Really Simple Syndication と名付けた。これにより RSS 0.91 に始まった RSS の混乱状態は収束することとなった。

コンテンツ配信指向及び RSS 0.91 からの互換性をふまえたためなのか、RSS 2.0 の RSS は”Really Simple Syndication”の略称となっている。

2003 年 7 月に、RSS 2.0 制定の中心人物である、デイヴ・ウィナー (Dave Winer) の移籍もあり、仕様もハーバード大学ロースクールのバークマンセンターに移管された。

#### 2.2.5 RDF(Resource Description Framework)

Web の目標である「マシンに理解可能な情報」の表現のためには、メタデータなどのリソース相互の関係を、特定のアプリケーションに依存しない形で叙示的に示す共通の方法が必要である。RDF は、主語 (リソース) と述語 (プロパティ)、そしてその目的語 (オブジェクト: プロパティの値) の三者関係によって、関係

の連鎖を辿ることができるようなデータモデルを記述する。

以下の図 2.1 に blog の例を示す。



図 2.1: Blog トップページの例 (<http://alejandro-rivero.ameblo.jp/>)

以下の図 2.2 に図 2.1 で示した、blog のRSS を示す。

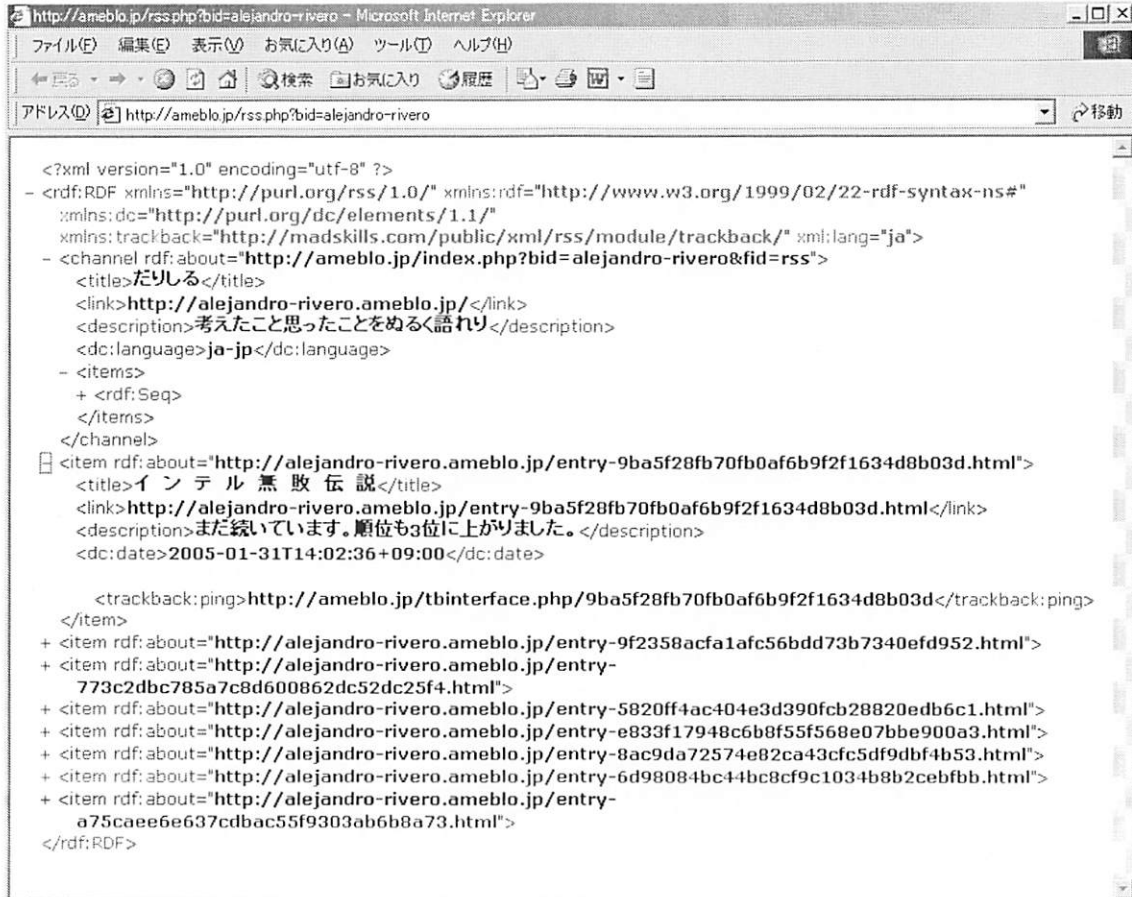


図 2.2: blog の XML 表示

## 2.3 XML

### 2.3.1 XML(eXtensible Markup Language)

XMLとは文書やデータの意味や構造を記述するためのマークアップ言語の一つであり、W3Cによって策定された。XMLの最初のバージョンであるXML1.0が1998年2月に勧告され、この改良版であるXML1.0Second Editionが2000年10月に勧告されている。XMLはHTMLと同じマークアップ言語だが、“extensible”とある通り、拡張できるという点で異なる。

Web上には無数の情報があり、これらは主にHTMLで記述されている。しかし、HTMLは元々論文記述用のマークアップ言語であり、タイトル・見出し・段落・強調等のいろいろな要素・属性を持っているが、それでも使用できる要素は限られている。たとえそれ以外にユーザが必要だと思う要素があっても、自由bに作る事はできず、そのため要素の内容を十分に表現することはできない。しかし、XMLでは、自由に要素や属性を作成・拡張でき、それによって情報の内容を適切に表現する事が可能になる。

特に情報を検索する場合には、HTMLには次のような限界がある。HTMLでマークアップされたリソースを検索する場合、指定するキーワードの相関関係を適切に設定する事は難しい。例えば、サーチエンジンを使って検索する場合、いくつかのキーワードを使って検索すると、全く関係の無い情報がヒットするという事があるが、これはサーチエンジンの不備というよりも、HTMLの限界であると言える。

XMLやHTMLは、元々出版業界で使うために文書記述言語としてISOが1986年に制定したSGMLから派生したものである。SGMLは汎用性が高いが、その半面、仕様が複雑で扱いが難しく十分に利用されていない。XMLはSGMLから多くの部分の仕様を受け継ぎつつも、今までにほとんど利用されなかった機能や一部の省略記法を取り除く事で仕様を簡略化し、特にインターネット上でのデータ交換を意識して設計されている。

このような特徴に加えて、XML文書はテキストデータであるため基本的にどんなシステムでも読み込む事ができる。またXMLの仕様は完全に公開されており、誰もが無料で自由に利用する事ができる。この「簡単に、誰もが、フリーに利用できる」という点も、大きなメリットである。

「XMLはHTMLに取って代わるものとなる」と言われる場合があるが、XMLはHTMLの代替になるというよりはむしろXHTML等のような言語を作成する言語、すなわちメタ言語として存在するものである。

XML文書を構成する要素は以下の三つに大別される。

\* XML 宣言

\* DTD (Document Type Definition)

\* XML インスタンス

XML宣言は、XMLの先頭に記述され、文字通りXMLである事を宣言するものである。ここでは“使用するXMLのバージョン”、“使用する文字セット”等を宣言する。また、XMLインスタンスとは、XMLの本体、すなわちタグでマークアップされたデータ群のことである。

DTDには個々の要素の定義等、そのXMLを記述するのに必要なものが全て定義されている。DTDがあるXMLをvalid XMLと言い、DTDが無いXMLをwell-formed XMLと言う。

DTDは必ずしも必要なものではないが、XMLでは例えば<商品>、<単価>、<数量>、<金額>等、要素・属性を自由に設定する事ができるので、DTDを記述する事によって、そのXML規則を他の場所でも使用する事ができるようになる。すなわち、複数サイトでのXMLの運用が可能になり、例えばWebショッピングモールを形成することができ、「複数の販売サイトから1番安い製品を検索する」というような事さえも可能となる。

このように記述されたXML文書が文法的に正しいかを検証し、表示するためにはXMLパーサが必要である。DTDがないXMLの場合は“開始タグと終了タグの対応”や“タグの入れ子関係”等の検証をし、DTDがある場合はそれ以外にそのDTDを読み込んで、“要素の並び順”や“要素の階層関係”等の整合性も合わせて検証する。

XMLはデータの構造を記述するものであり、従ってレイアウト情報を表すことができないので、XMLにはスタイル言語が用意されている。代表的なスタイル言語にはHTMLにも用いられているCSSと、XML文書をWebブラウザで見た場合の表現を記述するXSLがある。XSLは更に、XSLとXSLTに分類される。また、ハイパーリンク機能を実現するXLink/XPointerなどが用意されている [2]。

以下の表 2.1 に XML の例を示す。

表 2.1: XML の例

```
<?xml version="1.0" encoding="Shift_JIS"?>
<blog>
<title>XML@blog</title>
<Date>
  <Year>2005</Year>
  <Month>3</Month>
  <Day>14</Day>
</Date>
<comments>828</comments>
<trackback>56</trackback>
</blog>
```

### 2.3.2 SGML(Standard Generalized Mark-up Language)

文書の論理構造、意味構造を記述する言語のこと。タイトル、引用部分、著者など文書の中で特別な意味をもつ部分にマーク付けをすることができる。SGML オブジェクト間のリンクを記述するための XLink と、オブジェクトの参照方法を規定する XPointer で構成される。HTML のリンク機能に比べ大きく拡張されており、リンクをどう処理するかについての情報を記述したり、文書中のどの要素でもリンクを表現したり、1つのリンクに3つ以上のリンク先を定義したり、双方向のリンクを記述したり、コンテンツとリンクを切り離して管理したりといったことが可能である。

### 2.3.3 XSL(eXtensible Stylesheet Language)

XML 文書のスタイルシートを記述する言語。XSL の仕様はフォーマットリングオブジェクトと呼ばれる。XSL はページ書式を規定するオブジェクト、文書の段落、表、リストなどの様々な種類の組版用オブジェクトを定義する。つまり属性ごとで整理されたデータベースとなっている XML データを、視覚的に分かり易く並べ替えたり、デザイン性を加えて表示するための言語のこと。

#### 2.3.4 XSLT(XSL Transformations)

XMLによって記述された文書を他のXML文書に変換するための簡易言語。XML文書の構造を別の形式に変形するための変換ルールを記述するもので、記述されたXSLT文書は「スタイルシート」と呼ばれる。もともとはXSLの一部として変換処理を行なうために開発されたが、単独で使用することも可能である。おもに、XML文書からHTML文書やテキスト文書への変換などに使用される。XSLTによる変換を実行するためのソフトを「XSLTプロセッサ」と呼ぶ。

#### 2.3.5 XLink/XPointer(XML Linking Language/XML Pointer Language)

XML文書にハイパーリンク機能を持たせるための標準書式。オブジェクト間のリンクを記述するためのXLinkと、オブジェクトの参照方法を規定するXPointerで構成される。HTMLのリンク機能に比べると大きく拡張されており、リンクをどう処理するかについての情報を記述したり、文書中のどの要素でもリンクを表現したり、1つのリンクに3つ以上のリンク先を定義したり、双方向のリンクを記述したり、コンテンツとリンクを切り離して管理したりといったことが可能になっている。

## 第3章 blogのランキング手法

### 3.1 単一プロバイダ内のblogランキング手法

下記、livedoorBlog、gooBLOG、AmebaBlogの3つのblogサイトのblogランキングはそれぞれ独立している。

#### 3.1.1 livedoorBlogのblogランキング方法 (<http://blog.livedoor.com/>)

ページ閲覧数を表すページビューや投稿数をもとにして算出したポイントを用いてランキングを作成している。以下の図3.1にlivedoorBlogのトップページを示す。



図 3.1: livedoorBlog トップページ

### 3.1.2 gooBLOG の blog ランキング方法 (<http://blog.goo.ne.jp/>)

ページ閲覧数を表すPV(ページビュー)からランキングを作成している。以下の図??に gooBLOG のトップページを示す。

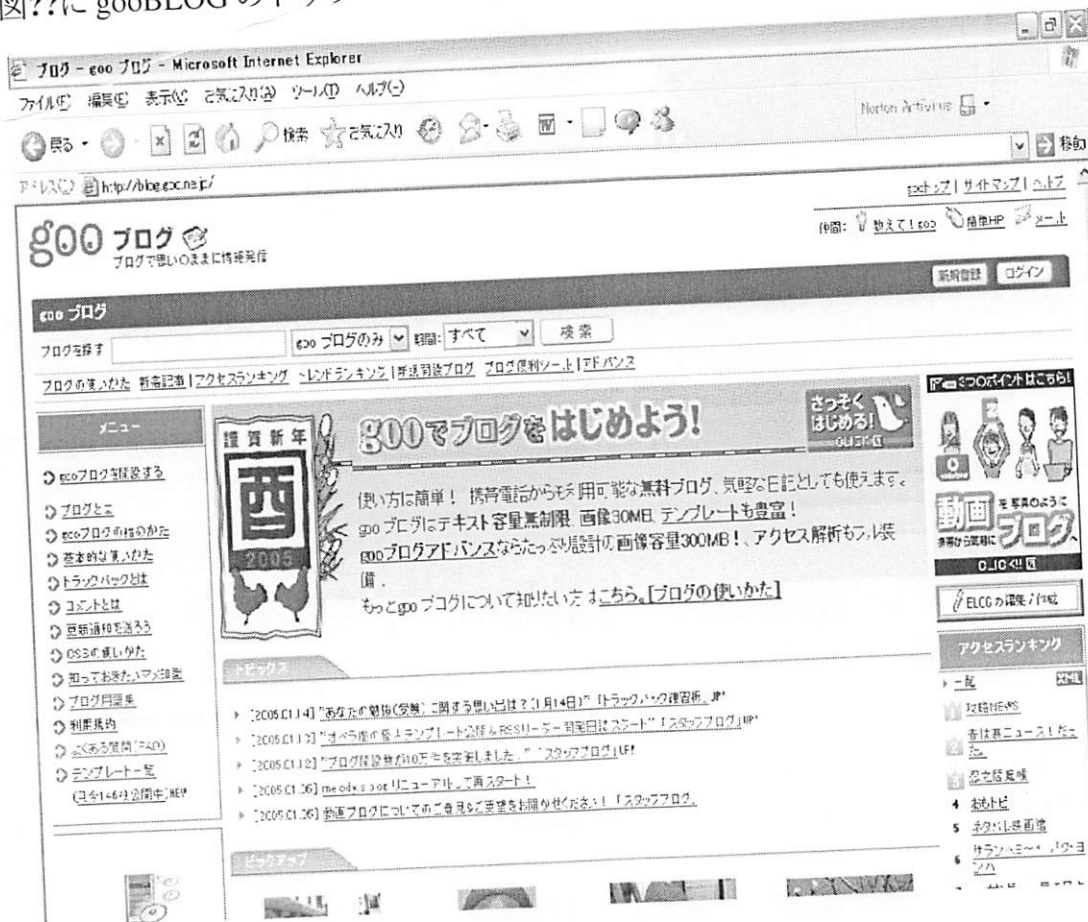


図 3.2: gooBlog トップページ

### 3.1.3 AmebaBlog の blog ランキング方法 (http://ameblo.jp/)

様々な数値をポイント化して独自の基準によって順位を付けている。コメントト  
ラックバックの数や更新の頻度をポイント化していると思われるが、基準の詳細  
については公開していない。以下の図 3.3 に AmebaBlog のトップページを示す。



図 3.3: AmebaBlog トップページ

上記、livedoorBlog、gooBLOG、AmebaBlog の3つの blog サイトの blog ランキ  
ングはそれぞれ独立している。

## 3.2 プロバイダに依存しないblogランキング手法

### 3.2.1 Blog Ranking(<http://blog.with2.net/>)

「Blog Ranking」に登録したblogから「Blog Ranking」のTOPページへのアクセス数(IN)と「Blog Ranking」のTOPページから登録したblogへのアクセス数(OUT)をリアルタイムで集計し、ランキングを作成している。IN/OUTともポイントの比重は同じ。

### 3.2.2 週刊ブログ王 (<http://www.doramix.com/rank/>)

「週刊ブログ王」のTOPページから「週刊ブログ王」に登録したblogへのアクセス数を集計してランキングを作成している。

## 第4章 Web ページ重要度を測る手法

### 4.1 PageRank

PageRank (ページランク) は Google 社が開発し、自社の検索エンジンに搭載している Web ページの重要度の判定技術、あるいは算出された各 Web ページの重要度の指数のことである。つまりそのページがどのくらい重要なのかを示すアルゴリズムのことである [3]。

「多くの良質なページからリンクされているページは、やはり良質なページである」という再帰的な関係をもとに、ページの重要度を計算している。あるページから別のページへのリンクをリンクされたページへの「支持投票」とみなし、それにリンク元のページの重要度(そのページの被リンク数)の重みをつけて加算し、投票数によりそのページの重要性を判断している。重要度の高いページにリンクを張られたページの評価は大幅に高くなるが、重要度の低いページにリンクを張られたページはあまり評価は上がらない [4]。つまり重要度の高いページによって投じられた票はより高く評価され「重要なページ」になっていく。こうした分析によって高評価を得た重要なページには高い PageRank (ページ順位) が与えられ、検索結果内の順位も高くなる。ページの内容は影響しない。

検索エンジンは従来、ページの内容と検索された単語との関連性の高さを判断し、検索結果の表示順位を決めていたが、Google 社は PageRank の導入により、どれだけ多くの人々が注目しているかという新しい指標を検索エンジンに持ち込んだ。同社ではページ内容の関連性の高さと PageRank の両方を総合的に判断して検索結果を表示している。ただし、どちらをどれくらい重視しているかは公表されていない。

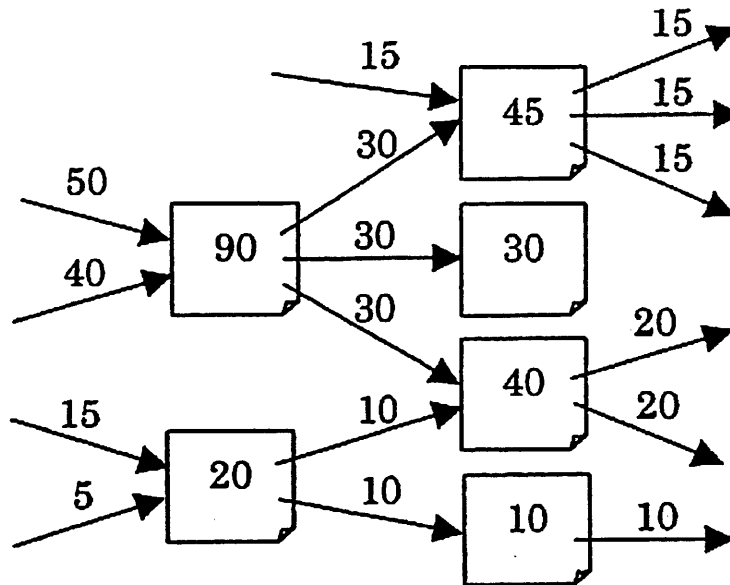


図 4.1: PageRank の概念図

#### 4.1.1 基本的な計算式

$u$  を Web ページ、 $F_u$  を  $u$  にリンクされているページの集合とする。また、 $N_u$  を  $u$  から出されているリンクの数 ( $N_u = |F_u|$ ) とし、 $c$  を一般化のための定数、そして  $u$  からリンクされているページ集合を  $B_u$  とする。このときページ  $u$  における pagerank の値  $R_{(u)}$  は以下の計算式によって計算される。

$$R_{(u)} = c \sum R_{(v)} / N_v (v \in B_u) \quad (4.1)$$

#### 4.1.2 行列計算による方法

全てのページに対して適応するために、行列計算を用いて計算を行う。各要素を以下に表した行列  $A$  を考える。もし  $u$  から  $v$  へとリンクが張られている場合には  $A_{u,v} = 1/A_u$ 、もしリンクが無い場合には  $A_{u,v} = 0$  となる。そして、 $R$  をそれぞれの Web ページの持つ pagerank の値を要素とするベクトルとした場合に

$$R = cAR \quad (4.2)$$

という式を得る。すなわち、 $R$ は行列  $A$  の固有ベクトルである。実際には  $R$  に対して  $A$  を繰り返し適応することによって値を計算することになる。

図 4.2 のようにリンクがつながっている Web 空間の場合を例にとって実際に行列を求めてみる。このようにリンクの関係が成り立っている場合、求める行列式  $A$  は式 (4.3) のようになる。

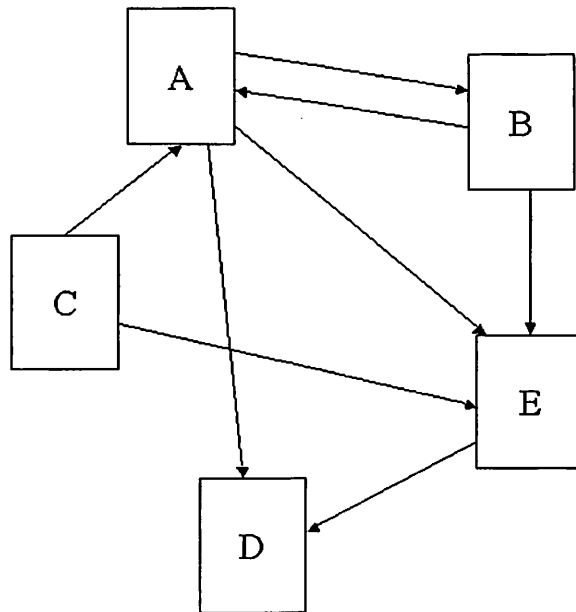


図 4.2: リンク例

$$\begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 1 \\ 1/3 & 1/2 & 1/2 & 0 & 0 \end{pmatrix} \quad (4.3)$$

#### 4.1.3 リンクを用いないアクセスの考慮

このアルゴリズムには問題が存在する。いくつかのページがリンクによってループしていたり、その数ページの間だけでリンクの関係が成り立っていたりする場合に、これらのページの値が無限大に発散してしまう点である。この問題を解決するために PageRank アルゴリズムは以下のように拡張されている。

「ユーザは、多くの場合は現在のページに存在するリンクをたどって移動するが、時々には全く無関係なページにジャンプする」というブラウジングモデルを考慮に入れた。「時々」というのは具体的には15%という数字に固定して計算している。ユーザは85%はリンクをたどるが、15%は突然無関係なページに飛んでいくということである。

$E_{(u)}$  を page jump を示すベクトルとした場合、式(4.1)とこのベクトルを用いて pagerank の値  $R'(u)$  は以下の計算式で計算が可能である。

$$R'_{(u)} = c \sum R'_{(v)} / N_v + c E_{(u)} (v \in B_u) \quad (4.4)$$

( $c$ は最大の値をとり、 $|R'| = 1$ を満たす)

この計算も同様に行列で計算が可能となる。行列  $E$  を全て行列要素が等しく(全ての要素が正である)、かつ  $|E|=0.15$  となる行列とする。また、この行列はリンクを用いないアクセスに対応する行列である。

この時、これらの行列は次の式を満たすことになる。

$$R' = c(A R' + E) (|R'| = 1) \quad (4.5)$$

ここで、 $|R'|=1$  であるため、1の全ての要素が1である行列として、上の式は以下の式に置き換えることができる。

$$R' = c(A + E \times 1) R' \quad (4.6)$$

この計算式も式(4.2)の計算と同様に  $(A + E \times 1)$  を繰り返し適用することによってベクトルを導き出すことが可能となる [5]。

## 4.2 HITS(Hypertext Induced Topic Search)

### 4.2.1 HITS アルゴリズム

適合度と重要度を両立させたランキング検索方法の一つである [6]。

ページの信頼性評価にあたり、まずはトピック（キーワード）が記述されているか、アンカーテキストが貼られているページが抽出される。次に抽出ページ群の中から、トピックのコミュニティ構造が解釈される。このとき、外部からの被リンクの多さに加えて、多くリンクされる Web ページに対する自リンクの多さも考慮している。これは、信頼性の高いページへのリンクの多いページは、ユーザにとって道標となる、という考えに基づく。そして、特定ジャンルのハブページ（優良リンク集）とハブページから良くリンクされるオーソリティページ（優良コンテンツ）が取り出される。さらにはオーソリティ同士のリンク構造も解析され、順位付けされる。HITS では、反復的な計算によって、Web グラフ中の各 Web ページのオーソリティスコアとハブスコアを計算し、オーソリティスコア順に Web ページを出力する。オーソリティ度の大きさが Web ページの最終的な信頼性評価となる。これにより、与えられたトピックに関連する Web ページ群の中から重要な Web ページが検索される [7]。

### 4.2.2 オーソリティ

特定トピックに関する情報源のことで、ハブとして価値の高いページからリンクを張られている。

### 4.2.3 ハブ

オーソリティへのハイパーリンクの集合のことで、オーソリティとして価値の高いページへリンクを張っている。

以下の図 4.3 にハブとオーソリティの概念図を示す。

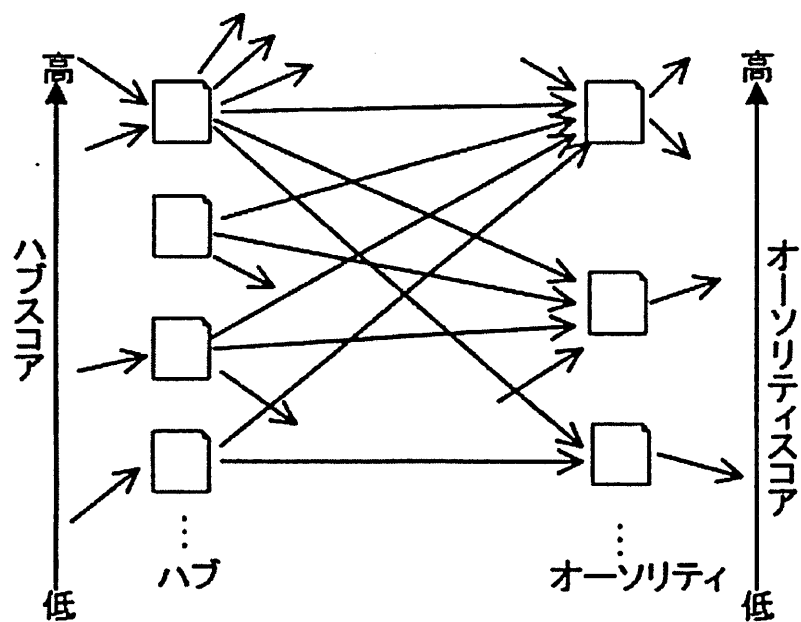


図 4.3: ハブとオーソリティの概念図

## 第5章 PageRankを利用したblog評価方法

本研究では Web ページ重要度を測るアルゴリズムのひとつである PageRank の手法を用いて blog 人気度の評価を行う。

PageRank は前述の通り多数引用されるページは信頼できる、また信頼できるページに引用されるページも信頼できるという再帰的な関係を仮定して Web ページの重要さを定義した指標をすものである。

PageRank の手法から、ページ A からページ B へのリンクが引用関係（ページ A がページ B を引用）を表すので、Web ページの被リンク数を求めることで引用度が計算できる。まずこの手法を blog 評価に適用して人気度の評価を行う対象の blog の被リンク数を求める。この情報は HTML ソースから得ることができないため、google Web service を利用して各 blog の被リンク数を求める。

次に、各々のリンクの重みとして、リンク元のページの引用度スコアに応じた値を用いる。この手法を blog 評価に適用し、blog に対してリンクを張ったページがどれほどの重要度を持っているかを調べるために人気度の評価を行う blog に対してリンクを張ったページの中の google 検索上位 10 件について、それぞれのページの被リンク数をさらに求める。

本研究では評価の対象とする blog の被リンク数、被リンクページの上位 10 件の被リンク数の合計数を足したものを blog 人気度の評価手法とする。

例として livedoorBlog ランキング 2 位の「社長日記」(<http://blog.livedoor.jp/takapon-jp/>)を用いて本手法を説明する。

(1) 評価対象である blog がどの程度の推薦を受けているかを表す被リンク数を求める。図 5.1 に google による被リンク数の計測の一例を示す。

(2)(1) で求めた被リンク数 1910 件の検索結果の上位 10 件のページについてそれぞれ被リンク数を求め、それらを足し合わせる。

(1) で求めた被リンク数 1910 と (2) で求めた被リンク数の合計 121 を足し合わせた 2031 が「社長日記」の blog 評価となる。



図 5.1: google による被リンク数の計測

## 第6章 実験

PageRank を用いた blog の評価方法がどれほど有効であるかを調べるために、既存のランキングと比較を行った。独立したランキングを持っているプロバイダの中から「livedoorBlog」、「gooBLOG」、「amebaBlog」の3つを、プロバイダに依存しない登録制の blog ランキングの中から「Blog Ranking」、「週刊ブログ王」の2つを合わせた合計5つのランキングを実験に使用する。

実験結果は以下のような項目に分かれる。

\*順位：既存のランキングの順位

\*blog名(URL)：評価を行う blog の名前と URL

\*評価：各プロバイダによる blog の評価

\*被 link 数：評価を行う blog の被リンク数

\*TOP10：評価を行う blog にリンクしているページの中の、google 検索上位 10 件の被リンク数の合計

以下に実験結果の表と、例として livedoorBlog ランキングとの比較のグラフを図 6.1 に示す。

表 6.1: livedoorBLOG - 1 (<http://www.livedoor.com/>)

順位	blog 名 (URL)	評価	被 link 数	TOP10
1	アキバ Blog( <a href="http://blog.livedoor.jp/geek/">http://blog.livedoor.jp/geek/</a> )	43889	1970	343
2	社長日記 ( <a href="http://blog.livedoor.jp/takapon.jp/">http://blog.livedoor.jp/takapon.jp/</a> )	21629	1910	121
3	ひろぶろ ( <a href="http://blog.livedoor.jp/stardom/">http://blog.livedoor.jp/stardom/</a> )	20669	504	51
4	RPG 版ウォ(瀉) チィ!( <a href="http://blog.livedoor.jp/rpgwatcher/">http://blog.livedoor.jp/rpgwatcher/</a> )	17111	115	115
5	HANBAI のデイトレ日記 ( <a href="http://daytrade.livedoor.biz/">http://daytrade.livedoor.biz/</a> )	15991	131	9
6	ザイーガ ( <a href="http://blog.livedoor.jp/parumo_zaeega/">http://blog.livedoor.jp/parumo_zaeega/</a> )	13061	1120	86
7	IT 企業のお仕事 ( <a href="http://blog.livedoor.jp/digital_tokyo/">http://blog.livedoor.jp/digital_tokyo/</a> )	10902	3	0
8	無料占いで開運して幸せを掴むブログ ( <a href="http://blog.livedoor.jp/urauranai/">http://blog.livedoor.jp/urauranai/</a> )	10410	15	32
9	なんでも評点 ( <a href="http://rate.livedoor.biz/">http://rate.livedoor.biz/</a> )	9785	148	287
10	うばーのお茶会 ( <a href="http://blog.livedoor.jp/uirou/">http://blog.livedoor.jp/uirou/</a> )	9646	364	261
11	週刊少年 Blog!!( <a href="http://blog.livedoor.jp/hitokata/">http://blog.livedoor.jp/hitokata/</a> )	9521	0	0
12	[Z]ZAPA ブログ ( <a href="http://blog.livedoor.jp/zapanet/">http://blog.livedoor.jp/zapanet/</a> )	9508	32	28
13	♪さきっちょ&はあちゅう♪の悪あが記 ( <a href="http://blog.livedoor.jp/firstlove2/">http://blog.livedoor.jp/firstlove2/</a> )	9456	371	47
14	プチ整形推進派：元祖しゃちょう日記 ( <a href="http://blog.livedoor.jp/hiroxi1492/">http://blog.livedoor.jp/hiroxi1492/</a> )	9431	1420	165
15	倉木麻衣の Mai.K Diary( <a href="http://kuraki.livedoor.jp/">http://kuraki.livedoor.jp/</a> )	9246	724	40
16	激安★超特価商店街 ( <a href="http://buy.livedoor.biz/">http://buy.livedoor.biz/</a> )	9190	224	28
17	小太郎ぶろぐ ( <a href="http://blog.livedoor.jp/kotaro269/">http://blog.livedoor.jp/kotaro269/</a> )	9167	491	43
18	ワラタニッキ ( <a href="http://blog.livedoor.jp/banmaco03/">http://blog.livedoor.jp/banmaco03/</a> )	8849	0	0
19	貧乏だけど心は萌え ( <a href="http://moe.livedoor.biz/">http://moe.livedoor.biz/</a> )	8843	272	53
20	まじかる☆めもりーず ( <a href="http://blog.livedoor.jp/magical_lounge/">http://blog.livedoor.jp/magical_lounge/</a> )	7768	0	0
21	りんご映画日記 ( <a href="http://blog.livedoor.jp/movie_ringo/">http://blog.livedoor.jp/movie_ringo/</a> )	7758	54	32
22	おれ日記 ( <a href="http://blog.livedoor.jp/orenikki/">http://blog.livedoor.jp/orenikki/</a> )	7622	549	61
23	さきっちょの就職日記 ( <a href="http://blog.livedoor.jp/sakiccho1228/">http://blog.livedoor.jp/sakiccho1228/</a> )	7585	845	384
24	0!特選ショップ ( <a href="http://www.moristar.com/">http://www.moristar.com/</a> )	7187	214	61
25	高橋がなりブログ 虎の声 - SOD( <a href="http://blog.livedoor.jp/sod/">http://blog.livedoor.jp/sod/</a> )	6750	349	168
26	峯田和伸の★朝焼けニャンニャン ( <a href="http://blog.livedoor.jp/mineta1/">http://blog.livedoor.jp/mineta1/</a> )	6428	294	12
27	キム兄のなで肩日記 ( <a href="http://blog.livedoor.jp/kim1963/">http://blog.livedoor.jp/kim1963/</a> )	6169	760	18
28	- go - ( <a href="http://monaken.livedoor.biz/">http://monaken.livedoor.biz/</a> )	5600	67	6
29	はあちゅう主義 ( <a href="http://blog.livedoor.jp/springflavor122/">http://blog.livedoor.jp/springflavor122/</a> )	5212	110	24
30	おもしろフラッシュ倉庫はっぴべ ( <a href="http://www.omosiroflash.com/">http://www.omosiroflash.com/</a> )	5113	8	1

表 6.2: livedoorBLOG - 2 (<http://www.livedoor.com/>)

順位	blog 名 (URL)	評価	被 link 数	TOP10
31	コマ斬り!( <a href="http://blog.livedoor.jp/lxl/">http://blog.livedoor.jp/lxl/</a> )	4762	5	41
32	特価 blog919!!( <a href="http://blog.919.bz/">http://blog.919.bz/</a> )	4504	184	6
33	livedoor Blog 開発日誌 ( <a href="http://blog.livedoor.jp/staff/">http://blog.livedoor.jp/staff/</a> )	4482	194	259
34	佐々木健介「嫁バカ日誌」( <a href="http://kensuke.livedoor.biz/">http://kensuke.livedoor.biz/</a> )	4350	223	27
35	*独身〇しまりえの超 節約日記* ( <a href="http://blog.livedoor.jp/marie_/">http://blog.livedoor.jp/marie_/</a> )	4313	53	33
36	大阪・日本橋でんでんタウン 電気街日誌 ( <a href="http://blog.livedoor.jp/ponbashi/">http://blog.livedoor.jp/ponbashi/</a> )	4212	113	71
37	蛆虫日誌@仮タイトル ( <a href="http://cybazzi.livedoor.biz/">http://cybazzi.livedoor.biz/</a> )	4151	0	0
38	イラストレーターのあきまん日記 ( <a href="http://blog.livedoor.jp/akiman7/">http://blog.livedoor.jp/akiman7/</a> )	4122	20	26
39	株で生活 D/Aの Speculation Life( <a href="http://blog.livedoor.jp/d_a/">http://blog.livedoor.jp/d_a/</a> )	3960	9	607
40	血の雨を降らす製氷機。( <a href="http://blog.livedoor.jp/gou_6431184/">http://blog.livedoor.jp/gou_6431184/</a> )	3932	9	1
41	アジアの真実 ( <a href="http://blog.livedoor.jp/lancer1/">http://blog.livedoor.jp/lancer1/</a> )	3558	3	0
42	忍者 RED( <a href="http://blog.livedoor.jp/ninja_red/">http://blog.livedoor.jp/ninja_red/</a> )	3504	360	56
43	モナぼ ( <a href="http://blog.livedoor.jp/monapopu/">http://blog.livedoor.jp/monapopu/</a> )	3500	60	51
44	ゲームショップ店長のお仕事とか ( <a href="http://gameshop.livedoor.biz/">http://gameshop.livedoor.biz/</a> )	3490	139	5
45	ina tekken( <a href="http://blog.livedoor.jp/inatekken/">http://blog.livedoor.jp/inatekken/</a> )	3446	37	66
46	俺暇 ( <a href="http://blog.livedoor.jp/gyroz/">http://blog.livedoor.jp/gyroz/</a> )	3332	612	158
47	Finding Lupo( <a href="http://blog.yamjun.com/">http://blog.yamjun.com/</a> )	2936	275	57
48	log( <a href="http://blog.livedoor.jp/adoruk626/">http://blog.livedoor.jp/adoruk626/</a> )	2918	354	23
49	韓国 CM を見よう！韓コマ!! ( <a href="http://blog.livedoor.jp/kancoma/">http://blog.livedoor.jp/kancoma/</a> )	2852	37	545
50	雅楽多 blog( <a href="http://gutti.livedoor.biz/">http://gutti.livedoor.biz/</a> )	2840	181	25
51	ルーの 4753 ライブドア株式投資 ( <a href="http://blog.livedoor.jp/piro2004/">http://blog.livedoor.jp/piro2004/</a> )	2766	133	16
52	えべっさんの商売繁盛ブログ ( <a href="http://ebessan.livedoor.biz/">http://ebessan.livedoor.biz/</a> )	2757	395	15
53	BOOKTERMINAL Blog( <a href="http://bookterminal.livedoor.biz/">http://bookterminal.livedoor.biz/</a> )	2739	0	0
54	eLeMeN - FINAL FANTASY XI DATABASE( <a href="http://blog.livedoor.jp/elemen/">http://blog.livedoor.jp/elemen/</a> )	2698	27	44
55	Daily DORA DORA( <a href="http://blog.livedoor.jp/dailydoradora/">http://blog.livedoor.jp/dailydoradora/</a> )	2667	3	3
56	特価情報サイト『よろづ屋』( <a href="http://blog.livedoor.jp/yorodu/">http://blog.livedoor.jp/yorodu/</a> )	2627	18	11
57	ドスバラ ( <a href="http://driver.dospara.co.jp/">http://driver.dospara.co.jp/</a> )	2597	301	21
58	明日からモテる男になる ( <a href="http://blog.livedoor.jp/ikebukuro/">http://blog.livedoor.jp/ikebukuro/</a> )	2546	119	26
59	氣志團現象最終章 ( <a href="http://kishidan.livedoor.jp/">http://kishidan.livedoor.jp/</a> )	2518	194	54
60	PLAYON RSS リーダー開発日記 ( <a href="http://blog.livedoor.jp/tomohip/">http://blog.livedoor.jp/tomohip/</a> )	2510	9	0

表 6.3: gooBLOG(<http://blog.goo.ne.jp/>)

順位	blog 名 (URL)	評価	被 link 数	TOP10
1	攻略 NEWS( <a href="http://blog.goo.ne.jp/gamekan/">http://blog.goo.ne.jp/gamekan/</a> )	6378	17	984
2	忍之悶魔帳 ( <a href="http://blog.goo.ne.jp/ipod_mini/">http://blog.goo.ne.jp/ipod_mini/</a> )	5196	45	550
3	昔は裏ニュース! だった。 ( <a href="http://blog.goo.ne.jp/luranews/">http://blog.goo.ne.jp/luranews/</a> )	4869	175	180
4	おもトビ ( <a href="http://blog.goo.ne.jp/trick12/">http://blog.goo.ne.jp/trick12/</a> )	2967	84	62
5	サランヘヨ * バク・ヨンハ ( <a href="http://blog.goo.ne.jp/rioha0624/">http://blog.goo.ne.jp/rioha0624/</a> )	2546	40	527
6	ネタバレ映画館 ( <a href="http://blog.goo.ne.jp/kossykossy/">http://blog.goo.ne.jp/kossykossy/</a> )	2229	56	74
7	風俗嬢の性の語り部屋 ( <a href="http://blog.goo.ne.jp/kyoukopsychology/">http://blog.goo.ne.jp/kyoukopsychology/</a> )	2148	25	600
8	小さなゲーム屋の日々( <a href="http://blog.goo.ne.jp/ryo11max/">http://blog.goo.ne.jp/ryo11max/</a> )	1999	45	54
9	28歳OL 独り暮らし失恋レポート ( <a href="http://blog.goo.ne.jp/momodesugananika/">http://blog.goo.ne.jp/momodesugananika/</a> )	1848	39	94
10	まっくんのひとりごと ( <a href="http://blog.goo.ne.jp/makkun35/">http://blog.goo.ne.jp/makkun35/</a> )	1638	51	94
11	リュシウォン 柳時元 ryusiwon LOVE( <a href="http://blog.goo.ne.jp/rurey55/">http://blog.goo.ne.jp/rurey55/</a> )	1319	4	1036
12	ちきうの気持ち ( <a href="http://blog.goo.ne.jp/thankyou7/">http://blog.goo.ne.jp/thankyou7/</a> )	1286	40	46
13	けーたいのひょうか ( <a href="http://blog.goo.ne.jp/hyoukaketai/">http://blog.goo.ne.jp/hyoukaketai/</a> )	1276	19	636
14	自然災害クルマエニ ( <a href="http://blog.goo.ne.jp/xxxobabaxxx/">http://blog.goo.ne.jp/xxxobabaxxx/</a> )	1170	11	58
15	某梨華ヲタ日記 ( <a href="http://blog.goo.ne.jp/rikagm/">http://blog.goo.ne.jp/rikagm/</a> )	1144	179	194
16	でじたるな部屋 ( <a href="http://blog.goo.ne.jp/digital_/">http://blog.goo.ne.jp/digital_/</a> )	1133	293	27
17	ゲーム屋店長奔走記 ( <a href="http://blog.goo.ne.jp/otake_famicom/">http://blog.goo.ne.jp/otake_famicom/</a> )	1116	61	550
18	Kento!BLOG( <a href="http://blog.goo.ne.jp/prowrestle/">http://blog.goo.ne.jp/prowrestle/</a> )	1100	126	8
19	ほっかも !! ( <a href="http://blog.goo.ne.jp/alb22_kitano/">http://blog.goo.ne.jp/alb22_kitano/</a> )	1066	5	44
20	殿下さま沸騰の日々( <a href="http://blog.goo.ne.jp/takkie0516/">http://blog.goo.ne.jp/takkie0516/</a> )	1054	207	70
21	□□お暇なら来てね。□□ ( <a href="http://blog.goo.ne.jp/kakokako3/">http://blog.goo.ne.jp/kakokako3/</a> )	1002	169	49
22	コンビニ・フリーク ( <a href="http://blog.goo.ne.jp/kazumucha/">http://blog.goo.ne.jp/kazumucha/</a> )	972	463	195
23	ブーケット津波情報総合 ( <a href="http://blog.goo.ne.jp/with_hope_1226/">http://blog.goo.ne.jp/with_hope_1226/</a> )	965	0	0
24	けいとのお気に入り日記 ( <a href="http://blog.goo.ne.jp/ma100hide/">http://blog.goo.ne.jp/ma100hide/</a> )	963	260	31
25	京の昼寝 ♪ ( <a href="http://blog.goo.ne.jp/cyaz/">http://blog.goo.ne.jp/cyaz/</a> )	954	116	771
26	カネシゲタカシの野球と漫画☆夢日記 ( <a href="http://blog.goo.ne.jp/mng-53/">http://blog.goo.ne.jp/mng-53/</a> )	909	361	117
27	蒼穹のぺうげおっと 5 5 ( <a href="http://blog.goo.ne.jp/55_peugeot/">http://blog.goo.ne.jp/55_peugeot/</a> )	859	44	585
28	CAT-walk DOG-run! ( <a href="http://blog.goo.ne.jp/happy4649miyako/">http://blog.goo.ne.jp/happy4649miyako/</a> )	839	132	538
29	ひらりんの映画ブログ ( <a href="http://blog.goo.ne.jp/hiralyn/">http://blog.goo.ne.jp/hiralyn/</a> )	827	4	23
30	俺が斬る! ( <a href="http://blog.goo.ne.jp/deadman44_2004/">http://blog.goo.ne.jp/deadman44_2004/</a> )	824	3	82

表 6.4: アメーバ BLOG(<http://ameblo.jp/>)

順位	blog 名 (URL)	被 link 数	TOP10
1	実録鬼嫁日記 ( <a href="http://yugure.ameblo.jp/">http://yugure.ameblo.jp/</a> )	4370	233
2	コロコロザイーガ ( <a href="http://korokorozaeega.ameblo.jp/">http://korokorozaeega.ameblo.jp/</a> )	842	102
3	語源 blog( <a href="http://yumo-p.ameblo.jp/">http://yumo-p.ameblo.jp/</a> )	2780	2052
4	大富豪の方程式 ( <a href="http://aircc.ameblo.jp/">http://aircc.ameblo.jp/</a> )	16200	749
5	お笑い@サブリッ! ( <a href="http://owaraisapuli.ameblo.jp/">http://owaraisapuli.ameblo.jp/</a> )	1390	189
6	*魔法のことば* ( <a href="http://happy-magic.ameblo.jp/">http://happy-magic.ameblo.jp/</a> )	2900	2763
7	関西外食総研 ( <a href="http://hassy.ameblo.jp/">http://hassy.ameblo.jp/</a> )	4870	470
8	渋谷ではたらく社長の blog( <a href="http://shibuya.ameblo.jp/">http://shibuya.ameblo.jp/</a> )	6450	1364
9	Numeblog( <a href="http://numeri.ameblo.jp/">http://numeri.ameblo.jp/</a> )	635	58
10	お礼状ブログ ( <a href="http://letters.ameblo.jp/">http://letters.ameblo.jp/</a> )	921	55
11	晴海 2 歳の闘病日記 ( <a href="http://toubyou.ameblo.jp/">http://toubyou.ameblo.jp/</a> )	1410	813
12	— 笑いの絶えない育児をしよう — ( <a href="http://monet.ameblo.jp/">http://monet.ameblo.jp/</a> )	891	2645
13	ほぼ日刊! 人生を変える魔法のコトバ ( <a href="http://skyroute.ameblo.jp/">http://skyroute.ameblo.jp/</a> )	1830	298
14	越後屋のコンサドーレ札幌とつれづれいろいろ ( <a href="http://consa.ameblo.jp/">http://consa.ameblo.jp/</a> )	320	68
15	◎実践! ネット ビジネス 奮闘記 ( <a href="http://webshopes.ameblo.jp/">http://webshopes.ameblo.jp/</a> )	1050	373
16	ドラクエ 8 攻略とかネタとかマンガとかゲームとか音楽とか色々。 ( <a href="http://draque.ameblo.jp/">http://draque.ameblo.jp/</a> )	397	62
17	☆ココロ開くよりアシ開くのが先やろ、やっば☆ ( <a href="http://mami24.ameblo.jp/">http://mami24.ameblo.jp/</a> )	483	93
18	はじめての NET 懸賞生活 ( <a href="http://howto.ameblo.jp/">http://howto.ameblo.jp/</a> )	2230	251
19	★映画館アルバイトー風呂敷研究中★ ( <a href="http://cuhasi.ameblo.jp/">http://cuhasi.ameblo.jp/</a> )	2850	552
20	モモの味は幸せの味 ( <a href="http://eien-ni-aisiteru.ameblo.jp/">http://eien-ni-aisiteru.ameblo.jp/</a> )	2900	722
21	CMニュース blog( <a href="http://cmm.ameblo.jp/">http://cmm.ameblo.jp/</a> )	791	387
22	ガンダムを語ろう ( <a href="http://g-seed.ameblo.jp/">http://g-seed.ameblo.jp/</a> )	342	557
23	★★かずくんままのマナー日記+α ( <a href="http://kazukilove.ameblo.jp/">http://kazukilove.ameblo.jp/</a> )	1590	573
24	ハッピーなオンナ ( <a href="http://girlfight.ameblo.jp/">http://girlfight.ameblo.jp/</a> )	761	81
25	★★★★★あなたは海外で笑う日本一です。 ( <a href="http://interhumor.ameblo.jp/">http://interhumor.ameblo.jp/</a> )	305	10
26	ぐ-すか-い-すか中国大冒険! ( <a href="http://huhehaote.ameblo.jp/">http://huhehaote.ameblo.jp/</a> )	965	337
27	▲▽ 愛欲の解毒場 ▽▲ ( <a href="http://love-poison.ameblo.jp/">http://love-poison.ameblo.jp/</a> )	914	54
28	夕刊マダムの悠々優待生活 ( <a href="http://yuukanmadam.ameblo.jp/">http://yuukanmadam.ameblo.jp/</a> )	5050	1213
29	イケナイ宝箱 ようこそ鬱の世界へ ( <a href="http://motogi.ameblo.jp/">http://motogi.ameblo.jp/</a> )	1670	401
30	とにかく新作映画を観て紹介する人柱的ブログ ( <a href="http://harinezmi.ameblo.jp/">http://harinezmi.ameblo.jp/</a> )	1410	566

表 6.5: 人気 blog ランキング (<http://blog.with2.net/>)

順位	blog 名 (URL)	被 link 数	TOP10
1	uwasa.tv 芸能界の噂話 ( <a href="http://www.uwasa.tv/">http://www.uwasa.tv/</a> )	348	11
2	韓流なんて知らない ( <a href="http://anti-korea.jugem.jp/">http://anti-korea.jugem.jp/</a> )	41	45
3	お笑い@サブリッ! ( <a href="http://owaraisapuli.ameblo.jp/">http://owaraisapuli.ameblo.jp/</a> )	1390	189
4	明日の注目銘柄 by 7777 ( <a href="http://www16.ocn.ne.jp/kabu7777/nikki.html">http://www16.ocn.ne.jp/kabu7777/nikki.html</a> )	0	0
5	発狂ニュース島 ( <a href="http://blog.livedoor.jp/kenngou777/">http://blog.livedoor.jp/kenngou777/</a> )	40	26
6	28歳 OL 独り暮らし失恋レポート ( <a href="http://blog.goo.ne.jp/momodesugananika/">http://blog.goo.ne.jp/momodesugananika/</a> )	39	94
7	毎日がキャバリア ( <a href="http://odekakewanko.cocolog-nifty.com/odekake/">http://odekakewanko.cocolog-nifty.com/odekake/</a> )	272	41
8	カメラと一緒にパリでお散歩 ( <a href="http://ariaparis.exblog.jp/">http://ariaparis.exblog.jp/</a> )	0	0
9	ファミリーセール情報探しネット ( <a href="http://blogs.dion.ne.jp/family/">http://blogs.dion.ne.jp/family/</a> )	3	0
10	アジアの真実 ( <a href="http://blog.livedoor.jp/lancer1/">http://blog.livedoor.jp/lancer1/</a> )	3	0
11	桜子の美健康ダイエット・ダイアリー ( <a href="http://bikennraku.cocolog-nifty.com/dietdiary/">http://bikennraku.cocolog-nifty.com/dietdiary/</a> )	181	66
12	実践!株主優待生活-我輩は株主である-( <a href="http://blog.kabulife.com/">http://blog.kabulife.com/</a> )	53	9
13	知識の泉 Haru's トリビア ( <a href="http://amor1029.exblog.jp/">http://amor1029.exblog.jp/</a> )	1150	62
14	面白いサイトを見つけたよ。 ( <a href="http://omosiroi.jp/">http://omosiroi.jp/</a> )	1270	209
15	おとなの恋草子 ( <a href="http://otona.livedoor.biz/">http://otona.livedoor.biz/</a> )	0	0
16	パチスロ北斗の拳の攻略法を自慢するブログ ( <a href="http://blog3.fc2.com/tabako300/">http://blog3.fc2.com/tabako300/</a> )	0	0
17	ネット内職で起業!金持ち兄さんの日記 ( <a href="http://blog.livedoor.jp/rich_navi/">http://blog.livedoor.jp/rich_navi/</a> )	0	0
18	ビバ★韓国 ( <a href="http://kankoku.jugem.jp/">http://kankoku.jugem.jp/</a> )	58	46
19	芸能界総合研究所 ( <a href="http://plaza.rakuten.co.jp/ewriweb">http://plaza.rakuten.co.jp/ewriweb</a> )	122	61
20	さきっちよの就職日記 ( <a href="http://blog.livedoor.jp/sakiccho1228/">http://blog.livedoor.jp/sakiccho1228/</a> )	845	384
21	日本がダイスキな外人の Blog( <a href="http://blog.livedoor.jp/delama/">http://blog.livedoor.jp/delama/</a> )	41	516
22	珍問屋 ( <a href="http://tindonya.com/">http://tindonya.com/</a> )	21	34
23	まあ待て、ブログを借りる前にここを読め。 ( <a href="http://ebn.livedoor.biz/">http://ebn.livedoor.biz/</a> )	414	32
24	株/新選組 ( <a href="http://kabu.main.jp/">http://kabu.main.jp/</a> )	77	99
25	Jumpi n×2 BLOG( <a href="http://blog.livedoor.jp/yoshi1214/">http://blog.livedoor.jp/yoshi1214/</a> )	0	0
26	★激安パソコン情報★ ( <a href="http://blog.livedoor.jp/gekipaso/">http://blog.livedoor.jp/gekipaso/</a> )	15	239
27	レミのイタリア Lovely ニュース ( <a href="http://italy.seesaa.net/">http://italy.seesaa.net/</a> )	0	0
28	お礼状ブログ ( <a href="http://letters.ameblo.jp/">http://letters.ameblo.jp/</a> )	921	60
29	ハッカーへの階段 ( <a href="http://jii3.no-blog.jp/hacker/">http://jii3.no-blog.jp/hacker/</a> )	173	99
30	彼はアタシの捌け口 ( <a href="http://hakeguchi.exblog.jp/">http://hakeguchi.exblog.jp/</a> )	0	0

表 6.6: 週刊 blog 王 (<http://www.doramix.com/rank/>)

順位	blog 名 (URL)	被 link 数	TOP10
1	☆OL 玲香のフォト日記☆ ( <a href="http://yaplog.jp/reirei_ken/">http://yaplog.jp/reirei_ken/</a> )	0	0
2	パチスロ北斗の拳の攻略法を自慢するブログ ( <a href="http://blog3.fc2.com/tabako300/">http://blog3.fc2.com/tabako300/</a> )	0	0
3	ひきこもり日記 ( <a href="http://blog.livedoor.jp/hikki_woods/">http://blog.livedoor.jp/hikki_woods/</a> )	7	9
4	REQUIEM OVERDRIVE( <a href="http://yaplog.jp/kakipee/">http://yaplog.jp/kakipee/</a> )	39	17
5	+♪*.O.O.。ていたらく日誌。・O.*♪+( <a href="http://yaplog.jp/seiteki">http://yaplog.jp/seiteki</a> )	0	0
6	ひろぶろ ( <a href="http://blog.livedoor.jp/stardom/">http://blog.livedoor.jp/stardom/</a> )	504	115
7	ポリポリ劇場 2 ( <a href="http://umaiboo.ameblo.jp/">http://umaiboo.ameblo.jp/</a> )	41	36
8	嫁には言えない…( <a href="http://lidlle.ameblo.jp/">http://lidlle.ameblo.jp/</a> )	272	60
9	まんが喫茶からお小遣い稼ぎ ( <a href="http://blog.livedoor.jp/solitude_of_citroen/">http://blog.livedoor.jp/solitude_of_citroen/</a> )	0	0
10	今日のおせん ( <a href="http://blog.goo.ne.jp/osen-com/">http://blog.goo.ne.jp/osen-com/</a> )	0	0
11	恋愛人。 ( <a href="http://yaplog.jp/ahy-034/">http://yaplog.jp/ahy-034/</a> )	8	36
12	写真集ダイジェスト ( <a href="http://photodigest.seesaa.net/">http://photodigest.seesaa.net/</a> )	8	36
13	hommania+blog( <a href="http://hommania.lolipop.jp/cgi/mt/">http://hommania.lolipop.jp/cgi/mt/</a> )	164	111
14	1000 万円！ ネット収入サイトになろう！ ( <a href="http://10000000yen.seesaa.net/">http://10000000yen.seesaa.net/</a> )	0	0
15	教えて会計 ( <a href="http://blog.livedoor.jp/keiri_nita/">http://blog.livedoor.jp/keiri_nita/</a> )	1	18
16	OL 生活たららん日記♪ ( <a href="http://blog.livedoor.jp/bonbon2004/">http://blog.livedoor.jp/bonbon2004/</a> )	0	0
17	アフィリエイトで月 30 万稼ぐサイト … にする！ ( <a href="http://yui1218.seesaa.net/">http://yui1218.seesaa.net/</a> )	0	0
18	魁ていない！ 男塾 ( <a href="http://www.ladiesmarket.net/blog/">http://www.ladiesmarket.net/blog/</a> )	0	0
19	寿司はサビ抜きで。 ( <a href="http://d.hatena.ne.jp/kmkskst/">http://d.hatena.ne.jp/kmkskst/</a> )	0	0
20	諸行無常的徒酒柿映船坤譚鼎譚阿機 ( <a href="http://blog.zaq.ne.jp/hil9000/">http://blog.zaq.ne.jp/hil9000/</a> )	0	0
21	☆まいとゆいのHP☆新人タレント活動中 ( <a href="http://plaza.rakuten.co.jp/maiuyui/">http://plaza.rakuten.co.jp/maiuyui/</a> )	0	0
22	Find the Way to…( <a href="http://po-po.cocolog-nifty.com/">http://po-po.cocolog-nifty.com/</a> )	1	830
23	くるとく News Blog( <a href="http://blog.livedoor.jp/kurutoku/">http://blog.livedoor.jp/kurutoku/</a> )	9	126
24	パチンコの話。スロットの話。 ( <a href="http://plaza.rakuten.co.jp/connective">http://plaza.rakuten.co.jp/connective</a> )	0	0
25	勝手な恋 最後の恋 未練な恋 ( <a href="http://www.doblog.com/weblog/myblog/30512">http://www.doblog.com/weblog/myblog/30512</a> )	1	3
26	まーん通信簿。 ( <a href="http://blog.livedoor.jp/oyamoney/">http://blog.livedoor.jp/oyamoney/</a> )	1	0
27	悪徳商法被害者の会 ( <a href="http://blog.livedoor.jp/stop_akutoku/">http://blog.livedoor.jp/stop_akutoku/</a> )	1	10
28	:: 素敵な女性への道のり :: ( <a href="http://yaplog.jp/saori0831/">http://yaplog.jp/saori0831/</a> )	10	34
29	Prep State examination Log( <a href="http://blog.goo.ne.jp/s02980734/">http://blog.goo.ne.jp/s02980734/</a> )	2	0
30	OL 侍。 ( <a href="http://lovefc.ameblo.jp/">http://lovefc.ameblo.jp/</a> )	911	116

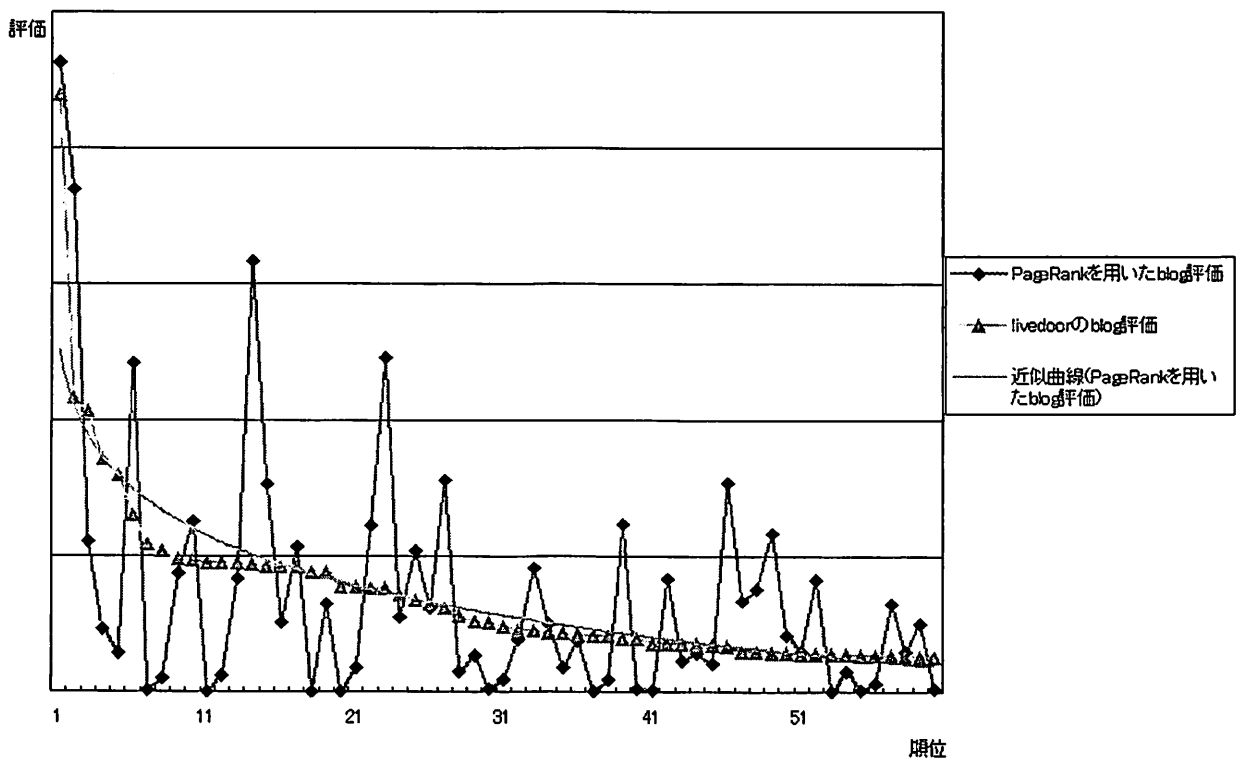


図 6.1: livedoorBlog ランキングの実験結果のグラフ

## 第7章 考察

本研究では blog の人気度を測る手法として、被リンク数と被リンクページの人気度に注目して blog 評価の実験を行ったが、各プロバイダによる blog 評価と本研究の blog 評価を比較すると多くの blog において評価が一致せず、おおまかな順位しか特定することができなかった。つまり、既存のランキングの主な blog 人気度の評価手法であるアクセス数と本研究の blog 人気度の評価手法である被リンク数は必ずしも比例関係にあるわけではないということがわかった。

livedoorBlog の実験結果のグラフを見ると、blog ランキング上位 60 件に入っているにも関わらず極端に被リンク数が少ないブログが存在することが分かる。特に 11 位、18 位、20 位、37 位、53 位の blog においてはその blog にリンクしているページが 1 つも存在しないという結果になった。この理由としては、本研究において被リンク数を調べるのに用いた google の backlink 機能に関係があると考えられる。

google ではリンク検索を使って被リンクページと被リンク数を求めることができるが、google の backlink 表示は当該ページへの全ての被リンクを表示するわけではなく I。リンク元のページが以前から登録されていたとしてもリンク元ページが被リンクページに及ぼす影響がごくわずかな場合には、backlink 一覧には表示されないという google リンク検索の仕様が存在する。以前は backlink 一覧には PageRank4 以下のページは表示されないといった明確な定義があったのだが、最近では PageRank が 2 や 3 でも表示されることがある。これは相対的に調査対象ページに対する各々の PageRank 貢献度が基準になっていると推測される。

プロバイダの blog ランキングでは上位に位置していて且つリンクを張っているページは存在するにも関わらず、google のリンク検索によって被リンクページが無いあるいは極端に少ないという結果になった blog は以上のような理由が考えられる。また今回は何万という blog の中から上位 60 件という狭い範囲での評価の比較だったので、細かい順位の判定が困難であったことも原因と考えられる。

## 第8章 おわりに

本研究の目的は、blog に対する評価の方法や評価基準の異なるプロバイダ間 blog に対して、統一ランキングを作る上で必要となる blog の人気評価手法の提案である。その手法として Web ページの重要度を測るアルゴリズムである PageRank を用いたが、各プロバイダの blog 評価と今回の手法による blog 評価では異なる箇所が多く出てしまった。そのため、新たに blog の人気度を測る blog 評価方法を考える必要があるが、blog 人気度の一番の目安となる各 blog のアクセス数はプロバイダによって公開されていない。そのため、第三者でも計測可能であり、且つ blog 評価を行うことができる手法でなければならない。

今後はランキング上位数十件という狭い範囲での blog 評価の比較であったので、より広い範囲で実験を行いより正確な blog 評価の比較を行うことと、今回は単一の手法でのみ blog の評価を行ったので、被リンク数の他に今回用いなかった blog の更新頻度や RSS フィードの購読者数、コメント、トラックバック数などを総合的に評価する方法が考えられる。

## 第9章 謝辞

本研究の遂行及び論文の作成において多大な御助言及び指導を賜った新納 浩幸 教官 (茨城大学工学部システム工学科) に深い感謝の意を表します。

また、御指導を頂きましたシステム工学科計算機応用学講座の教官の方々にも深く感謝します。

最後に、本研究を進めるにあたり助言、協力を頂きました岩崎 唯史教官 (茨城大学工学部システム工学科)、同研究室の紺野 憲一氏 (茨城大学大学院理工学研究科システム工学専攻)、藤井 文明氏 (茨城大学大学院理工学研究科システム工学専攻)、正木 裕一氏 (茨城大学大学院理工学研究科システム工学専攻)、谷津 哲平氏 (茨城大学大学院理工学研究科システム工学専攻)、大北 高広氏 (茨城大学システム工学科4回生)、木本 俊氏 (茨城大学システム工学科4回生)、高橋 宏直氏 (茨城大学システム工学科4回生)、藤村 元彦氏 (茨城大学システム工学科4回生) に深く感謝いたします。

## 参考文献

- [1] 山下清美：“ウェブログ・ネットワークング”  
[http://www.sendai-sentyuri.co.jp/sup/  
tsuushin7\(yamashita\).html](http://www.sendai-sentyuri.co.jp/sup/tsuushin7(yamashita).html)
- [2] 中山幹敏, 奥井康弘：“XML 完全解説 上/下”
- [3] “Google の人気の秘密”  
[http://www.google.co.jp/intl/ja/why\\_use.html](http://www.google.co.jp/intl/ja/why_use.html)
- [4] Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd：“The PageRank Citation Ranking: Bringing Order to the Web.1998”
- [5] 向享, 成凱, 上林彌彦：“利用履歴に基づく PageRank アルゴリズムの改良”
- [6] 阿部匡史, 豊田 正史, 喜連川 優：“アンカーテキストとリンク構造解析を用いた Web 情報検索の改善”  
[http://srd.yahoo.co.jp/PAGE=P/LOC=P/R=3/\\*-http://www.ieice.or.jp/  
iss/de/DEWS/proc/2003/papers/2-P/2-P-04.pdf](http://srd.yahoo.co.jp/PAGE=P/LOC=P/R=3/*-http://www.ieice.or.jp/iss/de/DEWS/proc/2003/papers/2-P/2-P-04.pdf)
- [7] J.Kleinber：“Authoritative Sources in a Hyperlinked Environment,Research ReportRJ10076(91892),IBM,1997”

## 付録A プログラムソースリスト

```
////////////////////////////////////  
//sotuken.java  
////////////////////////////////////
```

```
//被リンク数を求める。
```

```
import javax.swing.*;  
import java.awt.*;  
import java.awt.event.*;  
import java.net.*;  
import java.io.*;
```

```
/*public*/ class sotuken{
```

```
public static void main(String[] args){
```

```
//コンテナオブジェクト変数を宣言し、nullを代入
```

```
Container myContainer = null;
```

```
//JFrameクラスのインスタンスを作成
```

```
JFrame myFrame = new JFrame("被リンク数を数えるアプリケーション");
```

```

//レイアウトマネージャのインスタンスを作成
FlowLayout fLayout = new FlowLayout(FlowLayout.CENTER,3,30);

//コンテナ (フレーム) にレイアウトをセット
myFrame.getContentPane().setLayout(fLayout);

//JLabel クラスのインスタンスを作成
JLabel myLabel1 = new JLabel("URL");

//JTextField クラスのインスタンスを作成
final JTextField myText1 = new JTextField(50);

//JLabel クラスのインスタンスを作成
JLabel myLabel2 = new JLabel("URL 改");

//JTextField クラスのインスタンスを作成
final JTextField myText2 = new JTextField(50);

//JButton クラスのインスタンスを作成
JButton myButton1 = new JButton("実行");

//ボタンがクリックされた時のイベントを定義
myButton1.addActionListener(new ActionListener() {
public void actionPerformed(ActionEvent ae) {
    //処理
        String strInput = ("http://www.google.co.jp/search?as
+myText1.getText()+"&btnG=%8C%9F%8D%F5");
myText2.setText(strInput);

String browserName = "C:\\Program Files\\Internet Explorer\\
        iexplore.exe";

```

```
String url = strInput;
try{
Runtime.getRuntime().exec(new String[] {browserName, url});
}catch(IOException e) {
e.printStackTrace();
}
}
});
```

```
//JButtonクラスのインスタンスを作成
```

```
JButton myButton2 = new JButton("クリア");
```

```
//ボタンがクリックされた時のイベントを定義
```

```
myButton2.addActionListener(new ActionListener() {
public void actionPerformed(ActionEvent ae) {
```

```
//処理
```

```
myText1.setText("");
```

```
myText2.setText("");
```

```
}
```

```
});
```

```
//フレームにコンポーネントを追加
```

```
myFrame.getContentPane().add(myLabel1);
```

```
myFrame.getContentPane().add(myText1);
```

```
myFrame.getContentPane().add(myLabel2);
```

```
myFrame.getContentPane().add(myText2);
```

```
myFrame.getContentPane().add(myButton1);
```

```
myFrame.getContentPane().add(myButton2);
```

```

//フレーム (ウィンドウ) が閉じる際の処理を定義
myFrame.addWindowListener(new WindowAdapter() {
public void windowClosing(WindowEvent e) {
System.exit(0);
}
});

try {

//Look & Feel の設定

//windows スタイルの Look&Feel に設定
UIManager.setLookAndFeel("com.sun.java.swing.plaf.
                windows.WindowsLookAndFeel");

//設定を反映させる
SwingUtilities.updateComponentTreeUI(myFrame);

//エラー処理ブロック
} catch (Exception e) {
}

//フレームのサイズを設定して表示
myFrame.setSize(620,200);
myFrame.setVisible(true);
}
}

```

```
////////////////////////////////////  
//cyukan.c  
////////////////////////////////////
```

```
//コメントとトラックバックの数を数える
```

```
#include <stdio.h>  
#include <stdlib.h>  
#include <string.h>
```

```
#define LINESIZE 25600  
#define WORDSIZE 100
```

```
void quit(char *);
```

```
int main(int argc, char *argv[])  
{  
    FILE *f1;  
  
    char buf1[LINESIZE],  
          buf2[LINESIZE],  
          buf3[LINESIZE],  
          buf4[LINESIZE],  
          *s11 = ">Comments(", //livedoor  
          *s21 = ">TrackBack(",  
          *s12 = "#comments", //kokolog
```

```

*s22 = "#trackback",
*s13 = "comment_open(this.href); return false",//gooblog
*s23 = "trackback_open(this.href); return false",
*s14 = "</a></p>",//BIGLOBE
*s24 = "</a> / <a href=",
*cp1,*cq1,
*cp2,*cq2,
*cp3,*cq3,
*cp4,*cq4,
c1[3],t1[3],
c2[3],t2[3],
c3[3],t3[3],
c4[3],t4[3],cc4[3],tt4[3];

int r,a,b,
flag11,flag21,com1,coms1,tra1,tras1,count11,count21,
flag12,flag22,com2,coms2,tra2,tras2,count12,count22,
flag13,flag23,com3,coms3,tra3,tras3,count13,count23,
flag14,flag24,com4,coms4,tra4,tras4,count14,count24;

double COUNT11,COUNT21,COMS1,TRAS1,avecom1,avetra1,
COUNT12,COUNT22,COMS2,TRAS2,avecom2,avetra2,
COUNT13,COUNT23,COMS3,TRAS3,avecom3,avetra3,
COUNT14,COUNT24,COMS4,TRAS4,avecom4,avetra4;

if(argc != 2)quit("引数が違う prog datafile");

if((f1 = fopen(argv[1],"r")) == NULL)quit("ファイルが開けな
い");//1つめのキーワード

while(fgets(buf1,LINESIZE,f1) != NULL)

```

```

    {

    cp1 = strstr(buf1, s11);
    cq1 = strstr(buf1, s21);

if(cp1 != '\0')
    {
    flag11 = 1;

    a = cp1 - buf1;

    printf("\n\nコメントの数は%c",buf1[a+10]);

    c1[0] = buf1[a+10];
    c1[1] = '\0';

    if(buf1[a+11]!='')
    {
    c1[1] = buf1[a+11];
    printf("%c",buf1[a+11]);
    c1[2] = '\0';

    if(buf1[a+12]!=' ')
    {
    c1[2] = buf1[a+12];
    printf("%c",buf1[a+12]);
    }
    }

    com1 = atoi(c1);
    coms1 += com1;

```

```
count11++;
COMS1 = coms1;      //doubleに変換
COUNT11 = count11; //doubleに変換
avecom1 = COMS1/COUNT11;

}
```

```
if(cq1 != '\0')
{
    flag21 = 1;
    b = cq1 - buf1;

    printf("\nトラックバックの数は%c",buf1[b+11]);

    t1[0] = buf1[b+11];
    t1[1] = '\0';

    if(buf1[b+12]!='')
    {
        t1[1] = buf1[b+12];
        printf("%c",buf1[b+11]);
        t1[2] = '\0';

        if(buf1[b+13]!='')
        {
            t1[2] = buf1[b+13];
            printf("%c",buf1[b+13]);
        }
    }
}
```

```

    }

    tral = atoi(t1);
    tras1 += tral;
    count21++;
    TRAS1 = tras1;
    COUNT21 = count21;
    avetra1 = TRAS1/COUNT21;

    }
}

if(argc != 2)quit("引数が違う prog datafile");

/*      if((f1 = fopen(argv[1],"r")) == NULL)quit("ファイルが
開けない");//2 つめのキーワード

while(fgets(buf2,LINESIZE,f1) != NULL)
{

    cp2 = strstr(buf2, s12);
    cq2 = strstr(buf2, s22);

if(cp2 != '\0')
    {
    flag12 == 1;
    a = cp2 - buf2;

    if(buf2[a+11]!='C')
    {

```

```

printf("\n\nコメントの数は%c",buf2[a+39]);

c2[0] = buf2[a+39];
c2[1] = '\0';

if(buf2[a+11]!='')
{
c2[1] = buf2[a+11];
printf("%c",buf2[a+11]);
c2[2] = '\0';

if(buf2[a+12]!='')
{
c2[2] = buf2[a+12];
printf("%c",buf2[a+12]);
}
}

com2 = atoi(c2);
coms2 += com2;
count12++;
COMS2 = coms2; //doubleに変換
COUNT12 = count12; //doubleに変換
avecom2 = COMS2/COUNT12;

}
else
{break;}

}

```

```

if(cq2 != '\0')
    {
    flag22 == 1;
    b = cq2 - buf2;

    if(buf2[a+12]!='T')
    {
    printf("\nトラックバックの数は%c",buf2[b+11]);

    t2[0] = buf2[b+11];
    t2[1] = '\0';

    if(buf2[b+12]!='')
    {
    t2[1] = buf2[b+12];
    printf("%c",buf2[b+12]);
    t2[2] = '\0';

    if(buf2[b+13]!='')
    {
    t2[2] = buf2[b+13];
    printf("%c",buf2[b+13]);
    }

    }

    tra2 = atoi(t2);
    tras2 += tra2;
    count22++;
    TRAS2 = tras2;

```

```

COUNT22 = count22;
avetra2 = TRAS2/COUNT22;

    }
else
{break;}

}
}
    if((r = fclose(f1)) == -1)quit("ファイルが閉じれない");
*/
if((f1 = fopen(argv[1], "r")) == NULL)quit("ファイルが開けない");//3
つめのキーワード

while(fgets(buf3, LINESIZE, f1) != NULL)
{

    cp3 = strstr(buf3, s13);
    cq3 = strstr(buf3, s23);

if(cp3 != '\0')
    {
    flag13 = 1;
    a = cp3 - buf3;

    printf("\n\nコメントの数は%c", buf3[a+39]);

    c3[0] = buf3[a+39];
    c3[1] = '\0';

    if(buf3[a+40] != '<')

```

```

{
c3[1] = buf3[a+40];
printf("%c",buf3[a+40]);
c3[2] = '\\0';

if(buf3[a+41] != '<')
{
c3[2] = buf3[a+41];
printf("%c",buf3[a+41]);
}
}

com3 = atoi(c3);
coms3 += com3;
count13++;
COMS3 = coms3; //doubleに変換
COUNT13 = count13; //doubleに変換
avecom3 = COMS3/COUNT13;

}

```

```

if(cq3 != '\\0')
{
flag23 = 1;
b = cq3 - buf3;

printf("\\nトラックバックの数は%c",buf3[b+41]);

t3[0] = buf3[b+41];

```

```

t3[1] = '\0';

if(buf3[b+42]!='<')
{
t3[1] = buf3[b+42];
printf("%c",buf3[b+42]);
t3[2] = '\0';

if(buf3[b+43]!='<')
{
t3[2] = buf3[b+41];
printf("%c",buf3[b+41]);
}
}

tra3 = atoi(t3);
tras3 += tra3;
count23++;
TRAS3 = tras3;
COUNT23 = count23;
avetra3 = TRAS3/COUNT23;

}
}

if((r = fclose(f1)) == -1)quit("ファイルが閉じれない");

if((f1 = fopen(argv[1],"r")) == NULL)quit("ファイルが開けない");//4
つめのキーワード

```

```

while(fgets(buf4,LINESIZE,f1) != NULL)
    {
strcpy(c4,"");
strcpy(cc4,"");

cp4 = strstr(buf4, s14);
cq4 = strstr(buf4, s24);

if(cp4 != '\0')
    {
flag14 = 1;
a = cp4 - buf4;

c4[0] = buf4[a-1];
c4[1] = '\0';

if((buf4[a-1]=='0') || (buf4[a-1]=='1') || (buf4[a-1]=='2') ||
(buf4[a-1]=='3') || (buf4[a-1]=='4') || (buf4[a-1]=='5') ||
(buf4[a-1]=='6') || (buf4[a-1]=='7') || (buf4[a-1]=='8') ||
(buf4[a-1]=='9'))
    {

c4[0] = buf4[a-1];
c4[1] = '\0';

if((buf4[a-2]=='0') || (buf4[a-2]=='1') || (buf4[a-2]=='2') ||
(buf4[a-2]=='3') || (buf4[a-2]=='4') || (buf4[a-2]=='5') ||
(buf4[a-2]=='6') || (buf4[a-2]=='7') || (buf4[a-2]=='8') ||
(buf4[a-2]=='9'))

```

```

{
c4[1] = buf4[a-2];

c4[2] = '\0';

if((buf4[a-3]=='0') || (buf4[a-3]=='1') || (buf4[a-3]=='2') ||
    (buf4[a-3]=='3') || (buf4[a-3]=='4') || (buf4[a-3]=='5') ||
    (buf4[a-3]=='6') || (buf4[a-3]=='7') || (buf4[a-3]=='8') ||
    (buf4[a-3]=='9'))
{
c4[2] = buf4[a-3];
c4[3] = '\0';

}
}

if((buf4[a-3]!='0') && (buf4[a-3]!='1') && (buf4[a-3]!='2') &&
    (buf4[a-3]!='3') && (buf4[a-3]!='4') && (buf4[a-3]!='5') &&
    (buf4[a-3]!='6') && (buf4[a-3]!='7') && (buf4[a-3]!='8') &&
    (buf4[a-3]!='9'))
{
if((buf4[a-2]!='0') && (buf4[a-2]!='1') && (buf4[a-2]!='2') &&
    (buf4[a-2]!='3') && (buf4[a-2]!='4') && (buf4[a-2]!='5') &&
    (buf4[a-2]!='6') && (buf4[a-2]!='7') && (buf4[a-2]!='8') &&
    (buf4[a-2]!='9')) {
    cc4[0] = c4[0];
    cc4[1] = '\0';

}

else{cc4[0]=c4[1];
    cc4[1]=c4[0];}
cc4[2] = '\0';

```

```

    }
    else{cc4[0]=c4[2];
        cc4[1]=c4[1];
        cc4[2]=c4[0];

cc4[3] = '\0';
}

if((buf4[a-2]=='0')||(buf4[a-2]=='1')||(buf4[a-2]=='2')||
    (buf4[a-2]=='3')||(buf4[a-2]=='4')||(buf4[a-2]=='5')||
    (buf4[a-2]=='6')||(buf4[a-2]=='7')||(buf4[a-2]=='8')||
    (buf4[a-2]=='9'))
{
    cc4[0] = c4[1];
    cc4[1] = c4[0];
cc4[2] = '\0';

}

printf("\nコメントの数は%c%c%c",cc4[0],cc4[1],cc4[2]);

}

else{cc4[0]='\0';
    printf("\nコメントの数は%c",cc4[0]);}

com4 = atoi(cc4);
coms4 += com4;
count14++;
COMS4 = coms4; //doubleに変換
COUNT14 = count14; //doubleに変換
avecom4 = COMS4/COUNT14;

```

```

    }

if(cq4 != '\0')
    {
    flag24 = 1;
    b = cq4 - buf4;

    if((buf4[b-1]=='0') || (buf4[b-1]=='1') || (buf4[b-1]=='2') ||
        (buf4[b-1]=='3') || (buf4[b-1]=='4') || (buf4[b-1]=='5') ||
        (buf4[b-1]=='6') || (buf4[b-1]=='7') || (buf4[b-1]=='8') ||
        (buf4[b-1]=='9'))
        {

        t4[0] = buf4[b-1];
        t4[1] = '\0';

        if((buf4[b-2]=='0') || (buf4[b-2]=='1') || (buf4[b-2]=='2') ||
            (buf4[b-2]=='3') || (buf4[b-2]=='4') || (buf4[b-2]=='5') ||
            (buf4[b-2]=='6') || (buf4[b-2]=='7') || (buf4[b-2]=='8') ||
            (buf4[b-2]=='9'))
            {

            t4[1] = buf4[b-2];

            t4[2] = '\0';

            if((buf4[b-3]=='0') || (buf4[b-3]=='1') || (buf4[b-3]=='2') ||
                (buf4[b-3]=='3') || (buf4[b-3]=='4') || (buf4[b-3]=='5') ||
                (buf4[b-3]=='6') || (buf4[b-3]=='7') || (buf4[b-3]=='8') ||
                (buf4[b-3]=='9'))
                {

```

```
t4[2] = buf4[b-3];
```

```
}
```

```
}
```

```
if ((buf4[b-3] != '0') && (buf4[b-3] != '1') && (buf4[b-3] != '2') &&  
    (buf4[b-3] != '3') && (buf4[b-3] != '4') && (buf4[b-3] != '5') &&  
    (buf4[b-3] != '6') && (buf4[b-3] != '7') && (buf4[b-3] != '8') &&  
    (buf4[b-3] != '9'))
```

```
{
```

```
if ((buf4[b-2] != '0') && (buf4[b-2] != '1') && (buf4[b-2] != '2') &&  
    (buf4[b-2] != '3') && (buf4[b-2] != '4') && (buf4[b-2] != '5') &&  
    (buf4[b-2] != '6') && (buf4[b-2] != '7') &&  
    (buf4[b-2] != '8') && (buf4[b-2] != '9'))
```

```
    tt4[0] = t4[0];
```

```
else{tt4[0]=t4[1];
```

```
    tt4[1]=t4[0];}
```

```
}
```

```
else{tt4[0]=t4[2];
```

```
    tt4[1]=t4[1];
```

```
    tt4[2]=t4[0];}
```

```
if ((buf4[b-2] == '0') || (buf4[b-2] == '1') || (buf4[b-2] == '2') ||  
    (buf4[b-2] == '3') || (buf4[b-2] == '4') || (buf4[b-2] == '5') ||  
    (buf4[b-2] == '6') || (buf4[b-2] == '7') ||  
    (buf4[b-2] == '8') || (buf4[b-2] == '9'))
```

```
{
```

```
cc4[1] = c4[1];
```

```
cc4[2] = c4[0];
```

```

    }
    printf("\nトラックバックの数は%c%c%c\n", tt4[0], tt4[1], tt4[2]);

    }

    else{tt4[0]='0';
        printf("\nトラックバックの数は%c\n", tt4[0]);}

    tra4 = atoi(tt4);
    tras4 += tra4;
    count24++;
    TRAS4 = tras4;        //doubleに変換
    COUNT24 = count24;    //doubleに変換
    avetra4 = TRAS4/COUNT24;

    }
}

```

```

    if((flag11 == 1) || (flag21 == 1))
    {
        printf("\n\nコメントの平均値は%f です。", avecom1);
        printf("\nトラックバックの平均値は%f です。 \n", avetra1);
        printf("\nこのblogの評価は%f ポイントです。 \n", avecom1*
                                                    5+avetra1*25);
    }

```

```

/*
    if((flag12 == 1) || (flag22 == 1))
    {

```

```

printf("\n\nコメントの平均値は%f です。 ",avecom2);
printf("\n\nトラックバックの平均値は%f です。 \n",avetra2);
printf("\n\nこのblogの評価は%f ポイントです。 \n",avecom2*
                                             5+avetra2*25);
}

*/

if((flag13 == 1)|| (flag23 == 1))
{
printf("\n\nコメントの平均値は%f です。 ",avecom3);
printf("\n\nトラックバックの平均値は%f です。 \n",avetra3);
printf("\n\nこのblogの評価は%f ポイントです。 \n",avecom3*
                                             5+avetra3*25);
}

if((flag14 == 1)|| (flag24 == 1))
{
printf("\n\nコメントの平均値は%f です。 ",avecom4);
printf("\n\nトラックバックの平均値は%f です。 \n",avetra4);
printf("\n\nこのblogの評価は%f ポイントです。 \n",avecom4*
                                             5+avetra4*25);
}

}

void quit(char *s)
{
puts(s); exit(1);
}

```