

## 作成者責任の概念を用いた Web ページ評価法のサイトへの適用実験

絹川達也<sup>¶</sup> 平林真実<sup>†\*</sup> 清光英成<sup>\*</sup> 大月一弘<sup>\*</sup> 北村新三<sup>†</sup>

<sup>¶</sup>神戸大学大学院総合人間科学研究科<sup>†</sup>神戸大学大学院自然科学研究科  
<sup>\*</sup>岐阜県立国際情報科学芸術アカデミー<sup>\*</sup>神戸大学国際文化学部

Web ページを評価する方法は大きく分けて、1)内容を解析するもの、2)リンク構造を解析するもの、3)各ページからは収集できない外部情報 (URL などを利用してページ提供者についての情報など)を基準にするものがある。我々は、3)を組み合わせた 2)によるページ評価方法を提案してきた。本稿では外部情報の基準として web サイト内における「ページ情報に対する責任」を採用した場合のページ評価方法について詳しく検討する。また、提案方式による評価の妥当性や、他のリンク構造解析による評価との比較のための実験を行った。その結果は、提案方式の妥当性と、サイトにおける「責任」の評価、「責任」の所在に関する解析、に対する有効性を示している。

キーワード：WWW, ページランキング, 作成者責任

## Evaluation of a ranking method for pages of a web site using the latent accountability about authors and organizations

Tatsuya Kinugawa<sup>¶</sup> Masami Hirabayashi<sup>†\*</sup> Hidenari Kiyomitsu<sup>\*</sup>  
Kazuhiro Ohtsuki<sup>\*</sup> Shinzo Kitamura<sup>†</sup>

<sup>¶</sup>Graduate School of Cultural Studies and Human Science, Kobe University

<sup>†</sup>Graduate school of Natural Science, Kobe University

<sup>\*</sup>International Academy of Media Arts and Sciences

<sup>\*</sup>Department of Cross cultural studies, Kobe University

Analysis of the link structure is one of a major ranking method of Web pages recently. Nevertheless, web surfers have to choose the pages to read from the result of the method. The web surfers consider latent credibility of the authors and organizations that provides the pages. We propose a link-based page ranking method for reflecting this credibility. We consider the author's responsibility for the information of pages as a typical measure of credibility on a web site and evaluate the application of the method to a web site. The experimental result shows the validity of the assignment of our method for ranking the pages in a web site.

Key words: WWW, page ranking, page accountability

### 1 はじめに

Web における情報に対し検索技術を中心にいくつかのページ評価方法が検討されている [1]。キーワード検索や構文解析を基本とする内容に基づく方法から、最近では PageRank™ 方式 [2] や Kleinberg [3] による方法などに代表されるようなリンク構造

解析を基本とするものが注目されている。

一方、URL を見てページ作成者や作成組織を特定し、それをもとにページを評価する場合がある。これは、ページの内容やリンクの状態などの web 上には表れない、個人や組織に対する社会的な評価をページの信頼度評価に反映させていると考えることができる。ここでは、このページ評価につながる個

人の社会的信用を Social Credibility (SC) と呼ぶ。我々は、SC に基づく評価とリンク解析を組み合わせて web ページを評価する方法を提案している [4]。提案方式では、SC を利用した各ページの評価とリンクの意味を示す種類に応じたリンクによる評価を組み合わせることで、ページ評価を行っている。

各ページには SC 評価により初期点が与えられ、その上にリンクの種類に応じたページ評価を行うことで、従来は無視されていたリンクの意味を反映した、より正確なページ評価の実現を目標としている。リンクによる評価では、リンク元ページの値に応じた相対評価を行うこと、そのページに対する最も高い評価を採用する、という 2 つの方針を採用してモデル化を行っている。

本稿では、提案方式を明確な SC 評価の基準設定が可能なサイト内について適用を行った。サイト内では各ページ作成者が持つ web 上の管理領域をテリトリという単位で捕らえ、各テリトリに対して SC の値を記述する *SCI* (Social Credibility Index) を準備している。リンクの意味の記述のためにはリンク属性を記述している。

サイト内における本方式による評価は、一種の責任という概念との適合性が高いと考えられる。本論文では、サイトにおける作成者責任という概念に基づき、本方式をサイトに対して適用実験を行い、責任評価としての妥当性を示した。

## 2 ページ評価法

### 2.1 リンクによる評価

Web におけるリンク構造解析によるページ評価方法としては、Page [2] らによる PageRank™ 方式や Kleinberg [3] らによる hub/authority 法などがある。PageRank™ では、リンク構造から各ページにおける滞在確率を計算しスコアリングを行っている。また hub/authority 法では、検索語に対して関連性と人気度から得られる値を利用し hub ページ、authority ページを抽出している。これらの方法は [2] が Google™ [5] の基本方式として使われるようにページに対するひとつの有効な評価方法となっている。しかしながら、両方式とも「よくリンクされるページはよいページ」という仮定に基づいているため、「被リンクの少ない良いページ」の評価が低くなったり、「多数の友人間で相互にリンクしているペー

ジ」のようなページの評価が不用意に高くなったりするという問題もある。

### 2.2 社会的な信頼性による評価

検索サイトを用いたキーワード検索において多くの候補がリストアップされたとき、各ページの URL を見てそのページの作成者や組織を特定し中身を読むページを選択することは一般によく行われることである。これは社会的にその発信者がどのような立場にあるかの評価を暗黙のうちに利用していると考えられる。本稿ではこのような社会的な評価を「Social Credibility (SC) による評価」と呼ぶ。SC は、ページ内容や web 構造とは関係のない社会的な外部評価である。

別な言い方をすれば、このような評価は web ページの評価が web 上の情報のみで評価するだけでは限界があることに起因している。しかしながら、SC による評価だけでは、「社会的に知られていない人のよいページ」や「同一作成者や組織の作成したページの中からよいページ」を探すことは困難となる。

このような社会的な信頼性を示す外部情報を Social Credibility (SC) と呼ぶ。SC は組織や個人に対して社会的に与えられた作成情報に対する責任、信頼などに代表される評価基準とする。

SC は web においてもサイトや作成者といった社会的な単位に対する評価としては有効である。しかし、このページに対してこの評価を付与することは難しいという問題点がある。

### 2.3 提案方法

我々の研究では、リンク構造を利用した評価と SC を利用した評価を併せ持つページ評価方法を提案している。提案方法では、リンクによる評価と SC による評価を以下の方針によりページ評価に利用している。

#### ・SC による評価

SC が示す信頼性の指標を示すために Social Credibility Index (*SCI*) を用意し、これを利用し各ページに対する評価を行う。これによる社会的な外部評価をページ評価として反映させる。

#### ・リンク構造による評価

*SCI* を利用して評価されたページの評価は、リン

ク構造を利用した評価によって各ページの評価点となる。リンクによる評価では、リンクすることによる他者の評価を反映することができる。さらに、リンク作成の際にリンク元でリンクの意味を明示することでリンクの意図に応じた評価を行なうことができる。

評価においては、自ページの評価点に依存した相対評価、与えられた評価の最大値をそのページの評価点とする、という方法を採用している。

これらの方針に基づき、提案方法では各ページに図1のような点数付けが行われる。

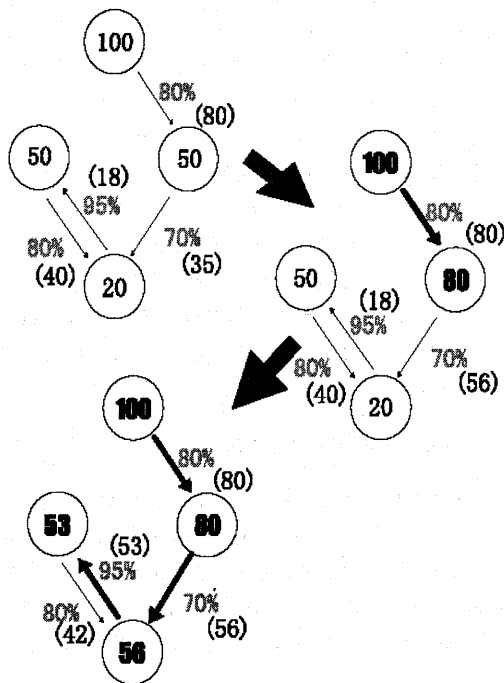


図1: 提案方法によるページ評価

この図では、円が各ページ、矢印がリンク、円内の数字は評価点、矢印に付いている数値(%)はリンクによる評価率と括弧内に評価値を示している。

①の状態では、まず初期点が各ページに与えられている。この状態での評価点はSCIを反映したものとなっている。ここでページaからページbへのリンクによる評価を反映すると②の状態となる。ページbの評価点が初期点である50点から80点に上が

っている。同様に、リンクによる評価を反映すると、③に示す状態となる。

③では、ページcへのリンクは2つ存在しているが、ページbは65点、ページcは40点としてページcを評価しているため、最大値を評価点とする方針からページbからのリンクが採用される。これによりページcの評価点は20点から65点へと上がることになる。さらにページcの評価が上がったため、ページcからページdへのリンクによる評価点(53点)がページdの初期点(50点)より高くなるため、ページcの評価によってページdの評価点が上がっている。

最終的に、ページaにSCIから与えられた評価点、ページb → ページc → ページdとリンクによる評価によって伝播していることになっている。このように、SCIによる点数を初期値として利用しながら、リンクによる相対評価が為されることで各ページに評価点が与えられることになる。

この方針に基づいて各ページの評価値 PaS(Page Score)を求めるための式は以下のように記述できる。

Webの構造を、各ページをノード、リンクをノード間の枝とする有向グラフGとすると、

$N$ : ノード数 (ページ数)

$P_i$ : ノード  $i$  ( $i=1,2,\dots,N$ )における PaSの値

$P_{base_i}$ : ノード  $i$ の初期 PaSの値

$A_{ij}$ : ノード  $i$ のノード  $j$ に対する評価

$p_{ij}$ : ノード  $i$ のノード  $j$ に与えられるスコア

$Src(i)$ : ノード  $i$ に対するリンクを持つノードの集合

$p_{ij}$ は相対的に与えられるので  $P_i$ と  $A_{ij}$ の関数

$$p_{ij} = f(P_i, A_{ij}) \quad (1)$$

と表す。

$P_i$ は以下の式で与えられる。

$$P_i = \text{Max}\{P_{base_i}, f(P_i, A_{ij})\} \text{ for all } i | n_i \in Src(j) \quad (2)$$

(2)式は  $P_i \geq p_{ij}$ を満たすならば付録に記したアルゴリズムを用いて  $O(N^2)$ の計算時間でとくことができる。

### 3 WWW サイトへの適用

#### 3.1 サイトにおけるページ作成者責任

2でのべたように、提案方法をWWWに対して適用し、

### 3.1 サイトにおけるページ作成者責任

2でのべたように、提案方法をWWWに対して適用し、PaSを求めるためには、

1. SCIの値の指定方法
2. PaSの初期値の付与方法
3. リンクによる評価点 $p_{ij}$ の付与方法

の3つを決定する必要がある。

本論文では、サイト内を対象として提案方式を適用した。サイト内においては、各ページ作成者の信頼性の記述が比較的容易であり、適切なSCIを決定できる。また、サイト内には推薦以外の意味をもつリンクが多数存在し、リンクの種類の違いによる評価の効果を発揮すると考えられる。

サイト内の社会的な信頼性の解釈として、我々はページ情報に対する責任という概念を用いている。ページ情報に対する責任とは、あるページに対して誰がどの程度の責任を保証しているかを示すものである。Webにおいては、ページを作成することだけでなく、別のページに対するリンクの作成にもある種の責任が発生すると考えられる。つまり、自分が作成したページの内容については作成者としての責任が発生し、リンクすることにより推薦を行う場合にもリンク先ページに対する一種の保証としての責任が発生すると考えることができる。このようなWebにおける責任という考えは、企業や大学、公共機関などの組織が自らの情報発信のために管理しているサイトに対して適用可能である。ページ情報に対する責任を示す評価としてPaSを解釈することで、本方式の方針は以下のように解釈できる。

- ・評価の最大値の採用  
最も責任を持つと宣言したものが責任を持つ。このことによりページに対する責任者が明確にすることができる。
- ・リンクによる相対評価  
リンク元ページの作成者の責任に応じてリンク先ページの責任を持つことができる。また相対評価により責任の伝播が発生する。

これらの解釈はページに対する責任として十分に

妥当なものであり、サイトにおける責任の指標としてPaSは有効なものであると考えている。

### 3.2 SCI

サイト内に適用範囲を限定した場合、SCはサイトを運営主体である組織が、各ページ作成者をどのように信用しているか、すなわちどの程度の責任を保障しているかを示すものと見ることができる。例えば、大学のサイトではたとえ同じ情報が書かれてあったとしても、学生が作成したものについては、大学として責任を負う事はできないが、公式情報として大学当局が作成したものには責任を負わなければならない。このような作成者単位で異なっている組織による責任の保証の度合いをSCIとして組織が指定することになる。

ここで、サイトにおいて各ページ作成者が持つ独自の管理領域をテリトリと呼ぶことにする。テリトリは、各ページ作成者がサイトにおいて自分のページを作成し、管理している領域であり、ページ作成者としての責任範囲を示している。ページ作成者の持つSCと作成したページ群を示すテリトリを明らかにすることにより、サイト内における責任を示すSCIを指定することができる。

### 3.3 PaSの初期値の付与方法

SCIを利用することで、サイトにおいてページ作成者が持つSCをページに対する責任として評価を行なうことができる。さらに各作成者はテリトリ内のページに対して自己評価を行なうことで、サイト内ページのPaSの初期値( $Pbase_i$ )を決定することができる。

まず、サイト側はSCIにより指定した各作成者のSCの値を作成者のテリトリトップページに対して $Pbase_i$ として与える。このトップページの $Pbase_i$ を基に、ページ作成者はリンクによってテリトリ内の他のページに対して相対的な評価を行なうことで、テリトリ内のページに評価が伝播する。

このようにサイトによるSCIを利用した評価が、各ページの $Pbase_i$ としてサイト内ページの評価として伝わることで、SCIによる各ページの間接的な評価として責任を示す値となる。

### 3.4 リンクによる評価点の付与方法

リンクによる責任の付与を行なう際には、リンクの持つ意味によりどのように責任として *PaS* の付与が行なわれるか異なってくる。例えば、責任の発生するリンクでは、テリトリ内自己評価での作成者責任を意味するリンク、テリトリ外への推薦を意味するリンク、そしてその責任の持ち方の度合いがあり、さらに責任の発生しないリンクなどが存在している。

責任の発生するリンクでは、テリトリ内では作成者として公式なページと個人的なページの区別、テリトリ外へのリンクでは責任を持てる推薦リンクと単なる参照リンクの区別などがある。また、テリトリトップページへの「戻る」を意味するリンクなどは、*SC* と無関係なナビゲーションや構造を示すためだけのリンクであり責任も発生しない。

我々は、このようなリンクの持つ責任の違いを  $A_{ij}$  において表す。  $P_i \geq p_{ij}$  を満たすように評価は自分のページの点数との比率で記述することにする。すなわち

$$f(P_i, A_{ij}) = P_i * V_{ij} \quad (3)$$

where  $0 \leq V_{ij} \leq 1$

また、 $A_{ij}$  の値は各ページ作成者がリンクを作成する際に直接ページに記述するものとする。

## 4 実験

実験には Linux 上の実験用 web サーバに神戸大学の国際文化学部のミラーを作成しこれを使用した。国際文化学部のサイトにおけるページ総数は 3055 またリンク数は 7734 であった。このサイトにおいては学生、教官、委員会などさまざまな作成者、組織が存在し、我々が考えているサイトの種類においては一般的なものである。適用に際し、以下に述べる方針で、*SCI* を記述するファイルとリンクの種類を示すものとしてリンク属性を用意した。なお、これらは記述の効率化や変更の容易さを実現し作成者の負担を減らすことを考慮したものとなっている。

### 4.1 サイトへの実装方法

#### *SCI*

*SCI* を作成する際にはクラスという単位によって実装を行なった。クラスとは組織における各ページ作

成者の立場によって分類したものである。クラスに分類することで、本来各ページ作成者毎に指定しなければならない *SC* の値を教員、学生などという単位で指定することができ、大幅な記述の簡略化が行なえる。*SCI* には各テリトリの URL、テリトリトップページの URL、そのテリトリのクラスを記載している。実験において使用したクラスを以下にあげる。

Faculty	…	学部としての公式な情報があるテリトリへのクラス
Division	…	それぞれの学科に関するテリトリへのクラス
Relevant	…	関連研究に関するテリトリへのクラス
Committee	…	学内委員会情報に関するテリトリへのクラス
Teacher	…	学部に所属する教官のテリトリへのクラス
Student	…	学部に所属する学生のテリトリへのクラス
Admin	…	サイトを実際に管理している部署のテリトリへのクラス
Project	…	学部でとり行っている研究に関するテリトリへのクラス

今回はそれぞれのクラスに対し以下のようなパラメータを設定した。

faculty	100	division	60
relevant	20	committee	80
teacher	50	student	30
admin	70	project	80

### リンク属性

3.4 において述べたリンクにおける責任の違いを示す種類 ( $A_{ij}$ ) を属性によって記述する。リンク属性は、ページ内の `<a href>` タグに作成者が直接書き込むものとする。リンク属性を利用することで、リンクの種類を値ではなく意味として記述することができ、属性の違い責任伝播量、サイト独自の責任伝播量などの調整を容易に行なえるようになっている。

我々が実験に適用する際に作成したリンク属性は以下のものである。

Equivalent: テリトリ内の責任の視点からはまったく同等なページへのリンク。各章ごとにページを設けた論文のページや、フレームの機能を利用している際に使用するとし、減衰はまったく行わないとした。

Official: テリトリ内のページに対し高い評価を付与する際のリンク。減衰を行うべきとは限定できないがここでは仮に 10% の減衰を行うことにした。この属性の減衰率による影響は後に実験結果により示す。

Personal: テリトリ内の責任をあまり付与しない際のリンク。但し、自分が作成した情報であるため多少は責任を負うべきと考えるため 60% の減衰を行うことにした。

Endorse: テリトリ外のページに対する強い推薦など、

高い評価を付与する際のリンク。実験では10%の減衰を行うことにした。

Introduce: テリトリ外のページで責任はもてないが紹介を行うようなものに対するリンク。リンクによる最低限の責任は負うものとして70%の減衰にとどめた。

Ignore: テリトリ内外を問わず PaS 伝播を行わないリンク。まったく点数の付与を行わない。

各属性による責任伝播率は次のようになる。

$$V_{ij} = \begin{cases} 1 & \text{if } A_{ij} = \text{equivalent} \\ 0.9 & \text{if } A_{ij} = \text{official} \\ 0.4 & \text{if } A_{ij} = \text{personal} \\ 0.9 & \text{if } A_{ij} = \text{endorse} \\ 0.3 & \text{if } A_{ij} = \text{introduce} \\ 0 & \text{if } A_{ij} = \text{ignore} \end{cases}$$

#### デフォルト処理による作業の省略

各作成者がすべてのリンクに対して責任の判断を行うことは実際には困難であることも考えられるために、リンクの属性がついていない場合にリンク先との関係性を判断し属性を自動で付与する処理についても検討した。

すでに述べたようなリンク属性はすべて必要であると考え、テリトリ内のページに対しては official を、テリトリ外のページに対しては introduce を、またすべてのテリトリトップに対しては ignore を付与する場合が多いという属性付与の傾向がある。実験に使用したプログラムにこの自動付与を組み込んだ。確かに厳密な評価を得るためにはすべてのリンクに対して各作成者が評価を行うことが必要であるがサイト側がいくつかのリンクに対して属性を付与するだけで残りは自動処理を行ったとしても妥当な結果は得られた。

#### 4.2 適用結果

図2は実装方法をもとに前述の SCI とリンク属性値を使用し、評価を行った結果と SCI のみを用いて点数を付与した結果との比較を行ったものである。

縦軸は各ページの点数を表し、横軸は各ページをあらわしており SCI の点数の高い順番にソートしたものである。

全体的に SCI だけでは表すことができなかった評価

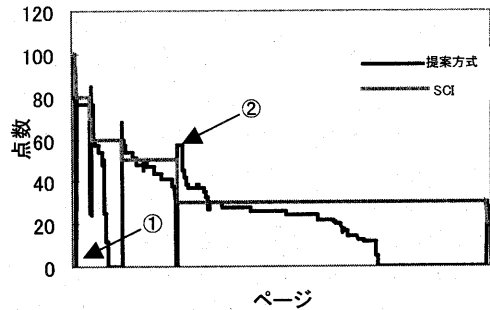


図2: 学部の全ページのスコア

が得られた。例えば①は学部が以前に開催したシンポジウムの開催案内のページである。このページはもともと学部のトップページからリンクが張られていたが、シンポジウムが終了したと同時にリンクが消されてしまったものである。このページの SCI は100点となるが、リンクによる評価の付与を受けていないため、実際の評価結果は0点となっている。また、阪神大震災に関するレポートをのせたページ②は学生が作成したものであるため SCI は30点であった。しかし、研究プロジェクトの一環であるとして、project から endorse され、58点となっている。

#### 4.3 他のリンク構造解析評価との比較

図3に PageRank™方式[2]との比較を示す。同方式の計算アルゴリズムとして、Haveliwala[6]の方法を用いた。比較を容易にするため、PageRank™による全ページの得点の合計点が、本方式の得点の合計と同じになるよう正規化した。

この結果より PageRank™による方式と SCI による評価の間に相関関係はないということがわかる。

表1は図3における①～⑦のページについて記したものである。

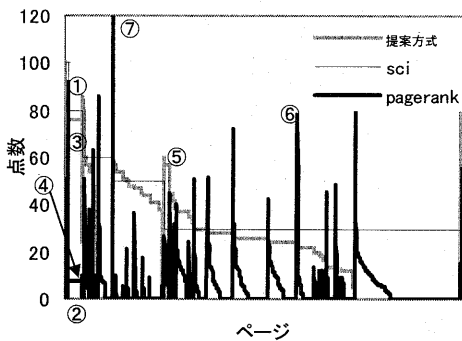


図3：提案方式とPageRank™方式の比較

ページの内容	提案方式	Page Rank™	SCI	責任の伝播経路
①学部のトップページ	100	91.7	100	100
②学部への交通情報	90	0.001	100	100→90→81
③EC委員会のトップ	80	69.9	80	80
④EC委員会の1ページ	76	7.9	80	80→76
⑤学生が作成した震災レポート	58	7.2	30	80→72→64.8→58.3
⑥ある学生のページ	28	85.5	30	30→27
⑦学部内研究者総覧のページ	80	600	50	80→80

表1：代表的ページのスコア

PageRank™において点数が高くなっているのは①③⑥⑦である。これらのページは他ページへのリンク、ならびに他ページからのリンクを多く含んでいる。特に⑥⑦はテリトリ内すなわち同一作成者の作成したページ間における相互リンクが多く作成されている。⑥の学生のページは自分の作成したページ間のみからリンクされており、他者からのリンクはない。PageRank™方式では被リンクが多いということは点数を上げる重要な要因であるため、このようなハブ的役目を持つトップページは点数が高くなる傾向が強い。しかし、特に責任などの基準に基づいているわけではないため、⑥のような学生が作成した私的な内容のページでも、①③⑦のような学部が作成したものでも無差別にリンク数が多ければ点数は高くなってしまっている。このような結果からページ責任を評価するのは難しい。

PageRank™では②④⑤のページの点数は①③⑥⑦

の点数に比べて著しく低い。②④は学部案内に含まれるページでリンクは少ないものの学部、あるいは公式委員会によって作成されたものである。これらのページは自テリトリのトップページから1ホップ、あるいは2ホップでリンクされているだけであり、このページからのリンクはない。このようなページは学部が大きな責任を持つためSCIが高いものの、リンク構造だけでは高い評価を与える事はできない。また⑤のように学生が作成したものではあるが学部公式プロジェクトのページから高い評価で推薦されているものもPageRank™では反映できていない。これに対して、本方式では②④に関しては、70点以上の高い評価、⑤に関してもそれらに準ずる評価となっている。

以上のような考察に基づくこの実験のようにサイトにおけるページ責任の概念など明確な評価基準を設けたページ評価を行う場合はSCIを利用した提案方式による評価のほうがPageRank™のようにリンク構造のみを用いた評価よりも適当であるといえる。

#### 4.4 責任の伝播

表1における「得点の伝播経路」はそのページを強く推薦したリンクのパス上のページにつけられた得点である。⑤のページの場合、プロジェクトのトップページ (PaS = 80) →同プロジェクトのサブページ (PaS = 72) →ある学生の震災研究のページ (PaS = 64.8) →レポートのページ (PaS = 58) という経路で得点が付与されている。すなわち、この学生のレポートのページの得点が58.3となったのは、プロジェクトのトップページから間接的に推薦されたことによる。58.3という評価を行ったのはそのページの作成者であるこの学生ではなくプロジェクトである。すなわちこのページに対し58.3という評価に対する責任を負っているのはプロジェクトであると考えられる。

ページ責任に対するこの考え方に基づけば、ページの得点の決定に用いられたリンクの経路はページに対する責任の経路を表すことになる。Webの構造グラフと責任の経路の関係を図4に示す。

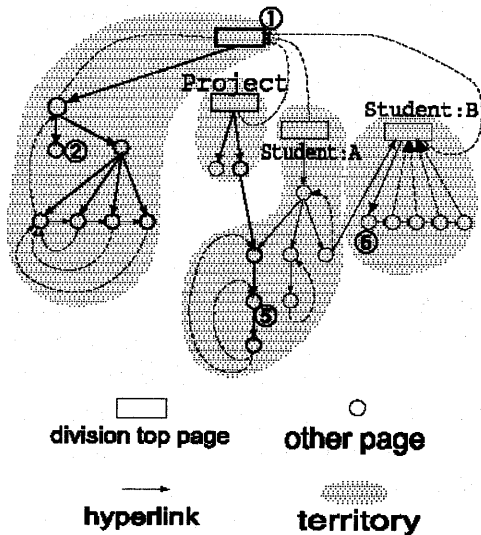


図4：責任の伝播ツリー

図において斜線部分がこのツリーの範囲をあらわす。図中の○はweb ページを、□は Territories のトップページをあらわす。また、太線はページの点数が伝播した経路を表す。この経路は Territories のトップページをルートとするツリーグラフになる。ツリーに含まれるページの評価に対する責任者(組織)はルートページの Territories 管理者となる。このツリーは図のように複数の Territories 間にもたががる。例えば、⑤は Project の責任伝播のツリーに含まれていることで点数が上がったことが明確に表現されている。逆に⑥は学生本人の責任ツリーに含まれているだけで、他からのリンクがなかったことか点数が上がらなかったことがわかる。

これによって各ページの他ページに対する責任保証の影響範囲を明確に把握することが可能となり、各ページにおいても誰の保証において、PaS が決定されたかを知ることができる。このような責任の伝播を表すツリーによって得られる構造は、ページ作成者のリンク作成時や、サイトの構成の編集時、SCI におけるパラメータの調整時などに非常に有効な判断材料となると考えられる。

#### 5 まとめ

本稿では web ページの評価法として SC とリンク構造を組み合わせた提案方法を web サイト内にペー

ジ責任の概念を用いて適用した。SC をリンク構造の評価における初期値とすることで web 上では得られない評価基準に基づいた評価が可能となった。責任の種類を表すためにリンクに属性を付与することで各ページの作成者はページごとに異なる自己評価を行うことができた。適用方法による結果はサイトとしての各ページに対するページ責任の評価を的確に表しており、評価としての有効性が示された。

#### 参考文献

- 1 佐藤光弘、梶浦正浩、多田智之、池野篤司、森口修 「WWW における情報検索技術の動向」 電子情報通信学会誌 vol.82 NO.12 1999 年 12 月, 1237-1242
- 2 Page, L., Brin, S., Motwani, R., Winograd, T., 'The PageRank Citation Ranking: Bringing Order to the Web', 1998
- 3 Kleinberg, J. M. Authoritative sources in a hyperlinked environment. In Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms 1998, 668 - 677
- 4 平林真実、堤智也、網川達也、大月一弘、北村新三 「大規模サイトにおける WWW ページ評価手法の検討」 情報学シンポジウム, 2001
- 5 <http://www.google.com/>
- 6 Taher H. Haveliwala, 'Efficient Computation of PageRank', Stanford Technical Report, 1999

#### 付録:Pj 導出アルゴリズム

すべての  $P_j$  は以下のアルゴリズムによって得ることができる。ここでは、 $PC_j$  を  $PaS$  の現在値を入れるもの、 $Label_j$  はノード  $j$  の現在の状態、 $cnode$  は現在注目するノードをそれぞれ示すものとする。

#### アルゴリズム

手順1: [初期化] すべてのノードに対して、 $PC_j = P_{base_j}$ ,  $label_j = 0$  とする。ここで、 $i=0$ ,  $cnode = 0$  とし、手順2に進む。

手順2: [最大の PaS を持つページへの注釈]  $Label$  の値が1でないすべてのノード  $k$  に対して、最大の  $PC_k$  を持つページ  $k$  を探す。[同値の場合、任意のノードを選択する]  $cnode = k$ ,  $P_k = PC_k$ ,  $label_k = 1$ ,  $i=i+1$  と置く。

$i > n$  ならば手順4に進み、それ以外なら手順3に進む。

手順3: [注目ノードからの PaS を計算する]

$cnode$  からのリンクを持ち、 $Label$  の値が1でないすべてのノードに対して、 $PC_j = \text{Max}(PC_j, P_{cnode} - V_{cnode,j})$  とする。手順2に進む。

手順4: 終了 [すべての  $P_k$  の値が計算された] □