

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5194778号
(P5194778)

(45) 発行日 平成25年5月8日(2013.5.8)

(24) 登録日 平成25年2月15日(2013.2.15)

(51) Int.Cl. F 1
G 0 6 F 1 7 / 3 0 (2 0 0 6 . 0 1) G 0 6 F 1 7 / 3 0 4 1 9 B

請求項の数 19 外国語出願 (全 11 頁)

| | | | |
|--------------|-------------------------------|-----------|--|
| (21) 出願番号 | 特願2007-330742 (P2007-330742) | (73) 特許権者 | 000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号 |
| (22) 出願日 | 平成19年12月21日(2007.12.21) | (74) 代理人 | 100070150 弁理士 伊東 忠彦 |
| (65) 公開番号 | 特開2008-217764 (P2008-217764A) | (72) 発明者 | ステリオス ステルギオウ アメリカ合衆国, カリフォルニア州 94305, スタンフォード, バーンズ・コート 74番 アパートメント308号 |
| (43) 公開日 | 平成20年9月18日(2008.9.18) | (72) 発明者 | ジャワハー ジャイン アメリカ合衆国, カリフォルニア州 95051, サンタ・クララ, リー・ドライブ 3575番 アパートメント10号 |
| 審査請求日 | 平成22年8月20日(2010.8.20) | | |
| (31) 優先権主張番号 | 60/876,546 | | |
| (32) 優先日 | 平成18年12月22日(2006.12.22) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 セッションベーススキエリのためのノードのランク付け

(57) 【特許請求の範囲】

【請求項1】

コンピュータが実行する方法であって、

セッションノードと前記セッションノードにリンクされた複数のリンクされたノードを含むノードのセットのモデルであって3以上の階層を含むモデルにアクセスする段階であって、前記リンクされたノードは親ノードと子ノードとを含み、親ノードは1つ以上の子ノードを前記セッションノードにリンクし、子ノードは前記子ノードを前記セッションノードにリンクする1つ以上の親ノードを有する、段階と、

前記セッションノードから所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットに対して確率分布を生成する段階であって、各子ノードはその親ノードの各々から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノードの各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する、段階と、

前記ノードのセットのランク付けに使用するため前記確率分布を伝える段階と、

を有し、

前記ノードのセットに対して前記確率分布を生成する段階は、一様な確率分布から始めて確率分布を反復的に計算する段階であって、生成される前記確率分布は反復的に計算された確率分布の収束の結果得られる、段階を含む、

方法。

【請求項2】

10

20

前記ノードはハイパーテキスト・マークアップ・ランゲージ (HTML) 文書である、請求項 1 に記載の方法。

【請求項 3】

前記所定数のリンクはユーザの我慢強さをモデル化したものである、請求項 1 に記載の方法。

【請求項 4】

前記所定数のリンクは 2 つのリンクである、請求項 1 に記載の方法。

【請求項 5】

前記所定割合は 1 / 2 である、請求項 1 に記載の方法。

【請求項 6】

前記ノードのセットのモデルはワールド・ワイド・ウェブの少なくとも一部をモデル化するグラフを含む、請求項 1 に記載の方法。

【請求項 7】

前記ランク付けは前記ノードのセットの検索を容易にして、ユーザからのクエリに対する応答を生成する、請求項 1 に記載の方法。

【請求項 8】

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

前記確率分布を平均化する段階と、

前記ノードのセットのランク付けに使用するため平均化した前記確率分布を伝える段階と、をさらに有する、請求項 1 に記載の方法。

【請求項 9】

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

ユーザにより指定された嗜好に基づき前記ノードのセットのランク付けに使用する前記複数の確率分布を伝える段階と、をさらに有する、請求項 1 に記載の方法。

【請求項 10】

セッションノードと前記セッションノードにリンクされた複数のリンクされたノードとを含むノードのセットのモデルであって 3 以上の階層を含むモデルにアクセスする段階であって、前記リンクされたノードは親ノードと子ノードとを含み、親ノードは 1 つ以上の子ノードを前記セッションノードにリンクし、子ノードは前記子ノードを前記セッションノードにリンクする 1 つ以上の親ノードを有する、段階と、

前記セッションノードから所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットに対して確率分布を生成する段階であって、各子ノードはその親ノードの各々から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノードの各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する段階、と、

前記ノードのセットのランク付けに使用するため前記確率分布を伝える段階と、

をコンピュータに実行させるプログラムであって、

前記ノードのセットに対して前記確率分布を生成する段階は、一様な確率分布から始めて確率分布を反復的に計算する段階であって、生成される前記確率分布は反復的に計算された確率分布の収束の結果得られる、段階を含む、

プログラム。

【請求項 11】

前記ノードはハイパーテキスト・マークアップ・ランゲージ (HTML) 文書である、請求項 10 に記載の プログラム。

【請求項 12】

前記所定数のリンクはユーザの我慢強さをモデル化したものである、請求項 10 に記載の プログラム。

【請求項 13】

10

20

30

40

50

前記所定数のリンクは2つのリンクである、請求項10に記載のプログラム。

【請求項14】

前記所定割合は1/2である、請求項10に記載のプログラム。

【請求項15】

前記ノードのセットのモデルはワールド・ワイド・ウェブの少なくとも一部をモデル化するグラフを含む、請求項10に記載のプログラム。

【請求項16】

前記ランク付けは前記ノードのセットの検索を容易にして、ユーザからのクエリに対する応答を生成する、請求項10に記載のプログラム。

【請求項17】

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

前記確率分布を平均化する段階と、

前記ノードのセットのランク付けに使用するため平均化した前記確率分布を伝える段階と、

をさらにコンピュータに実行させる請求項10に記載のプログラム。

【請求項18】

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

ユーザにより指定された嗜好に基づき前記ノードのセットのランク付けに使用する前記複数の確率分布を伝える段階と、

をさらにコンピュータに実行させる請求項10に記載のプログラム。

【請求項19】

セッションノードと前記セッションノードにリンクされた複数のリンクされたノードを含むノードのセットのモデルであって3以上の階層を含むモデルにアクセスする手段であって、前記リンクされたノードは親ノードと子ノードとを含み、親ノードは1つ以上の子ノードを前記セッションノードにリンクし、子ノードは前記子ノードを前記セッションノードにリンクする1つ以上の親ノードを有する、手段と、

前記セッションノードから所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットに対して確率分布を生成する手段であって、各子ノードはその親ノードの各々から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノードの各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する、手段と、

前記ノードのセットのランク付けに使用するため前記確率分布を伝える手段と、

を有し、

前記ノードのセットに対して前記確率分布を生成する手段は、一様な確率分布から始めて確率分布を反復的に計算する手段であって、生成される前記確率分布は反復的に計算された確率分布の収束の結果得られる、手段を含む、

システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はワールド・ワイド・ウェブ中のウェブページ等の相互リンクが張られたオブジェクトの検索に関する。

【関連出願】

この出願は、米国仮出願第60/876546号(2006年12月22日出願)の米国特許法第119条(e)の利益を主張するものである。

【背景技術】

【0002】

10

20

30

40

50

ウェブベース検索のためのノードのランク付けは重要な技術である。例えば、PAGERANKはグーグル社が使用しているものであるが、ワールド・ワイド・ウェブの構造を調べてウェブページをランク付けしている。しかし、PAGERANKは、成功しているが、ユーザの一般的な検索行動を正確にモデル化していない。

【発明を実施するための最良の形態】

【0003】

実施形態により、例えばワールド・ワイド・ウェブのウェブページ等の相互リンクが張られた文書またはその他のオブジェクトのコレクション (collection) を検索するためのセッションベースのユーザモデルを提供する。一部の実施形態では、典型的なユーザ行動を従来技術よりよく追跡 (track) できる、オブジェクトのセット全体の順序 (すなわちランク) を計算する。

10

【0004】

PAGERANKはウェブページをその間のリンクを調べてランク付けするアルゴリズムである。各ウェブページ P_i に付されるランクは、ユーザが、ランダムに選択したウェブページから始まるリンクをランダムに選択した場合に、いずれは P_i を訪れる (visit) 確率に対応する。ワールド・ワイド・ウェブを「サーフィン (surfing)」するこのモデルは、ランダムサーファーマデルである。

【0005】

ウェブに n ページ P_1, \dots, P_n あり、各ページ i は外への l_i 個のリンクを有すると仮定する。 $E_{ij} = 1$ であるとき、ページ i からページ j へのリンクがあることを示す。リンクが無ければ $E_{ij} = 0$ である。PAGERANKは n ページの各々のランクのベクトル R を計算する。ここで、 $R(i)$ はページ i のランクである。PAGERANKはこの計算を反復し、計算したベクトルが同じ (similar) になるまでこの計算を反復する。このアルゴリズムでは、次の計算を行う：

20

【数1】

$$R_0(j) = \frac{1}{n}$$

$$R_{t+1}(j) = \alpha \sum_i E_{ij} \frac{R_t(i)}{l_i} + (1-\alpha) \frac{1}{n}$$

30

PAGERANKでは、 α の値は 0.85 である。

【0006】

最初、ユーザがあるウェブページを訪れる確率はすべてのウェブページで等しい。ユーザが t 番目の反復時にウェブページ i をおとずれる確率は $R_t(i)$ である。ユーザはページ i 中の l_i 個のリンクの1つをランダムに選択するので、ユーザがウェブページ j にジャンプする確率は

【数2】

$$E_{ij} \frac{R_t(i)}{l_i}$$

40

である。PAGERANKは、この確率を、ユーザが次のステップにウェブページ j にいる確率 $R_{t+1}(j)$ に加える。ユーザがウェブページ j 以外のウェブページにジャンプする確率は $1 - \alpha$ であり、この確率は比較的小さい。

【0007】

ランダムサーファーマデルは分析が比較的容易であるが、現実のユーザが検索をするやり方を正しく捉えていない。対照的に、本発明の実施形態は検索セッションの概念に基づくモデルを利用する。

50

【 0 0 0 8 】

図 1 は、ウェブページのセットにわたってのユーザの検索行動例を示す図である。ユーザはセッションページ i から検索を開始すると仮定する。ユーザは外への l_i 個のリンクの 1 つから他のページ j に進み、検索をやめるか、セッションページ i に戻って異なるリンク経路に進むまで、リンクのシーケンスの行き来を続ける。例えば、ユーザはセッションページ i から検索を開始する。ユーザはノード 1 を訪れ、ノード 2 に進む。ここで、必要に応じて、ノードとはウェブページやその他のリンク可能オブジェクトを含むものとし、その逆の場合も同じとする。ユーザは、欲するものが見つからなければ、セッションページ i に戻り、他のリンクを選択し、ノード 3 まで行く。ユーザは、この方向性は望ましくないと判断し、再びセッションページ i に戻る。ユーザは、ノード 4 に行くリンクに進み、最終的にノード 5 に行く。

10

【 0 0 0 9 】

この検索方法は現実のユーザの典型的な行動により近いものである。ユーザは、長いリンク経路 (long path of links) をたどってページに行き着くのではなく、もっと小さいリンクシーケンスをたどり、その途中で (at the end) そのシーケンスをさらに先に進むかどうか判断する。ユーザは、そのシーケンスをさらに先に進まない場合、元のページに戻り他の経路に進む。かかる検索パターンはインテリジェントサーファーマデル (Intelligent Surfer Model) の基礎となる。一部の実施形態では、かかる検索パターンを使用し、ウェブページやその他のリンク可能オブジェクト (linkable objects) のランクを計算する。

20

【 0 0 1 0 】

H_i^d をセッションページ i から d 個のリンクを隔てたページのセットとする。 H_i をセッションページ i から最大 k 個のリンクを隔てたページのセットとする。一部の実施形態では、以下の一様分布から始めて、確率ベクトル R を反復的に計算する。

【 数 3 】

$$R_0(j) = \frac{1}{n}$$

$$R_{t+1}(j) = \alpha \sum_i C_{ij} R_t(i) + (1 - \alpha) \frac{1}{n}$$

30

R_0 は、すべてのノードが最初にセッションノードとして等しい確率を有することを意味している。 C_{ij} は、反復 t においてセッションページ i がノード j に貢献する確率 $R_t(i)$ の割合を示す。 C_{ij} は、ノード i とノード j の間に最大 k 個のリンクを有する経路がある場合のみ非ゼロとなる。一部の実施形態では、 C_{ij} は t とは独立なので、 C_{ij} を事前に計算する。

【 0 0 1 1 】

一部の実施形態では、 H_i 中のノードに確率 $R_t(i)$ を階層化して分配する。例えば、第 1 のステップですべての $R_t(i)$ を H_i^1 中の l_i 個のノードに分配する。簡単のため、一部の実施形態では、この分配は実質的に一様であると仮定する。第 2 のステップで、 j からリンクされた H_i^2 中のノードに、 H_i^1 中の各ノード j に分配された確率のパーセンテージを分配する。再度、簡単のため、一部の実施形態では、 j の確率の $1/2$ を j がリンクする l_j 個のノードに均等に分配するものとしてもよい。一部の実施形態では、このプロセスを k 回のステップに対して反復する。

40

【 0 0 1 2 】

図 2 乃至図 4 は、一部の実施形態で $R_t(i)$ を H_i 中のノードにいかにか分配するかを例示する分配例を示す。図 2 は一例であり、第 1 のステップで確率 $R_t(i)$ をノード v_1 、 v_2 及び v_3 に均等に分配する。ノード v_1 、 v_2 及び v_3 はノード i に集まっているものである。図 3 に例として示したように、これ以降のステップで、ノード v_1 、 v_2 及び v_3 の確率をそれぞれノード $[v_4, v_5, v_6]$ 、 $[v_6, v_7, v_8]$ 及び $[v_9$

50

、 v_{10}]に分配する。図4は の分配結果を示している。一部の実施形態では、この分析からファクタ C_{ij} を直接求め、一部の実施形態では、ファクタ C_{ij} は とは独立である。限定ではなく一例として、一部の実施形態では

【数4】

$$C_{iv_7} = \frac{1}{9} \text{ and } C_{iv_3} = \frac{1}{6}$$

である。一部の実施形態では、対応するファクタ C_{ij} をルックアップ (look up) することにより、各反復の動作を実行する。かかる実施形態では、各反復においてリンクを進む (link traversals) 必要はない。

10

【0013】

一部の実施形態では、このアルゴリズムの実行にはファクタ k が重要である。ファクタ k は、所望のウェブページを見つけるためにユーザが進んでもよいと思うリンクを示す。従って、ファクタ k はユーザの辛抱強さをモデル化したものである。一部の実施形態では、 $k=1$ の場合はPAGERANKの計算に戻る。従って、一部の実施形態はPAGERANKを包含している。

【0014】

図5は、セッションベースクエリのためにノードのランク付けをするシステム10の例を示す図である。システム10は、端末12とサーバ14とを含み、これらはネットワーク16を介して通信する。一部の実施形態では、ネットワーク16は、ローカルエリアネットワーク (LAN)、ワイヤレスLAN (WLAN)、ワイドエリアネットワーク (WAN)、メトロポリタンエリアネットワーク (MAN)、インターネットの一部、その他のネットワーク16、またはこれらの組合せである。本発明では、ネットワーク16は任意の適切なものでよい。リンク18は各端末12またはサーバ14をネットワークに結合する。一部の実施形態では、各リンク18は有線リンク、無線リンクまたは光リンクを含む。一部の実施形態では、各リンク18はLAN、WLAN、WAN、MAN、インターネットの一部、公衆交換電話網 (PSTN)、アクセスネットワーク、その他のリンクまたはこれらの組合せである。本発明では、任意の適切な端末12と任意の適切なサーバ14を任意の適切なネットワーク16に結合する任意の適切なリンク18を予定している。

20

30

【0015】

端末12によりユーザはネットワーク16を介して通信できる。限定ではなく一例として、端末12はコンピュータシステム (ノートブックまたはデスクトップコンピュータシステム)、携帯電話 (ワールド・ワイド・ウェブをブラウズする機能を含むものであってもよい)、その他の端末12を含む。サーバ14は、端末12にサーバ14の機能またはデータへのアクセスを提供する。限定ではなく一例として、サーバ14はウェブサーバを含んでもよい。ウェブサーバは端末12からのHTTP (Hyper Text Transfer Protocol) リクエストを受け取り、端末12にHTTPレスポンス及びウェブページ等の要求データを伝送する。ウェブページはHTML文書を含んでいてもよい。他の例として、サーバ14はアプリケーションサーバを含んでいてもよい。本発明では、任意の適切なサーバ14を予定している。

40

【0016】

一部の実施形態では、システム10の1つ以上のサーバ14は、1つ以上のウェブ検索エンジン20、1つ以上のランキングエンジン22、またはその両方を含む。一部の実施形態では、バスまたはその他のワイヤにより、1つ以上のウェブ検索エンジン20は通信のため1つ以上のランキングエンジン22に結合している。ウェブ検索エンジン20の例には、BAIDU、グーグル、LIVE SEARCH、ヤフー!検索などがある。本発明では、任意の適切なウェブ検索エンジン20を予定している。さらに、本発明は任意の適切な検索エンジンを予定しているが、これは必ずしもウェブ検索エンジン20である必

50

要はない。一部の実施形態では、ランキングエンジン 22 は、上記のように、ウェブページまたはその他のリンク可能オブジェクトをランク付けするハードウェア、ソフトウェア、または組み込みロジックコンポーネント (embedded logic component) またはこれらの 2 つ以上の組合せを含む。限定ではなく一例として、ランキングエンジン 22 はサーバ 14 のすべてのウェブページまたはその一部をランク付けし、1 つ以上のウェブ検索エンジン 20 はランキングエンジン 22 が生成したランキングを使って、端末 12 のユーザからの検索クエリに回答して検索結果を生成し、それを端末 12 のユーザに返す。

【0017】

一部の実施形態では、ユーザがセッションページ P から検索を開始して、そこから始まる一定数 N のリンクをたどるというユーザモデルを仮定してインテリジェントサーファーマodel化を試みる。ある確率で、ユーザは検索がうまく行かず、セッションページ P に戻って別のリンク経路を進む。一部の実施形態では、PAGERANK はかかるモデルの特別な場合 (N = 1 の場合) である。一部の実施形態では、検索結果を生成して返すために、ウェブページまたはその他の相互リンクされたオブジェクトをよりよくランク付けできる。

【0018】

図 6 は、セッションベースクエリのためにノードのランク付けをする方法例を示す図である。G = (V, E) はワールド・ワイド・ウェブをモデル化したグラフとする。ノード V はウェブページを表し、ノード間をつなぐエッジはウェブページ間のウェブリンクを表す。ユーザが最初のウェブページから進み、あきらめて最初のウェブページから再び開始する連続したリンクの数を N とする。数 N はユーザの我慢強さを表す。G と N が与えられると、アルゴリズムにより、インテリジェントサーファーマodelをたどるユーザに対応するノード V にわたる確率分布 D を求める。すなわち、インテリジェントサーファーマodelをたどるユーザは、確率分布 D において高い確率を有するノードに行く可能性が高い。本方法はステップ 100 で始まり、アルゴリズムにより V において一様な分布 $D_0 = U$ を生成する。ステップ 102 において、アルゴリズムにより、最終分布 D に収束するまで、反復的に「よりよい」分布 D_i を生成する。各反復 i において、アルゴリズムにより、v から N 個のエッジ以内で到達できるすべてのノードに対して、各ノード v に確率 $D_i(v)$ を分布させる。こうして分布 D_{i+1} を求める。一部の実施形態では、アルゴリズムにより、確率分布 D_i を U で重み付けして分布 D_{i+1} を生成する。ステップ 104 において、 D_{i+1} と D_i が互いに、例えば所定の 1 つ以上の基準に基づき非常に近くなるとアルゴリズムは終了する。ステップ 106 において、結果として得られた、ノードをランク付けする確率分布 D を伝えて、本方法が終了する。一部の実施形態では、必要に応じて、N の複数の値に対してアルゴリズムを複数回実行する。かかる実施形態では、結果として得られた分布を平均するか、またはユーザが検索を実行するときに選択肢を与える。

【0019】

本発明は、ここに説明した実施形態に対する、当業者が想到するであろうすべての変更、置換、変形、代替、修正を含むものである。同様に、適切な場合には、特許請求の範囲は、ここに説明した実施形態に対する、当業者が想到するであろうすべての変更、置換、変形、代替、修正を含むものである。

【0020】

なお、本発明の一部の実施形態を整理すると以下の通りである。

(付記 1) セッションノードと前記セッションノードにリンクされた複数のリンクされたノードとを含むノードのセットのモデルにアクセスし、前記リンクされたノードは親ノードと子ノードとを含み、親ノードは 1 つ以上の子ノードを前記セッションノードにリンクし、子ノードは前記子ノードを前記セッションノードにリンクする 1 つ以上の親ノードを有する段階と、

前記ノードのセットに対して、前記セッションノードから所定リンク数内のすべてのリンクされたノードに確率を分配する確率分布を生成し、各子ノードはその親ノードの各々から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノード

10

20

30

40

50

の各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する段階と、
前記ノードのセットのランク付けに使用するため前記確率分布を伝える段階とを有する
方法。

(付記 2) 前記ノードはハイパーテキスト・マークアップ・ランゲージ (HTML) 文
書である、付記 1 に記載の方法。

(付記 3) 前記所定数のリンクはユーザの我慢強さをモデル化したものである、付記 1
に記載の方法。

(付記 4) 前記所定数のリンクは 2 つのリンクである、付記 1 に記載の方法。

(付記 5) 前記所定割合は $1/2$ である、付記 1 に記載の方法。

(付記 6) 前記ノードのセットのモデルはワールド・ワイド・ウェブの少なくとも一部
をモデル化するグラフを含む、付記 1 に記載の方法。

10

(付記 7) 前記ランク付けは前記ノードのセットの検索を容易にして、ユーザからのク
エリに対する応答を生成する、付記 1 に記載の方法。

(付記 8) 前記セッションノードから相異なる所定リンク数内のすべてのリンクされた
ノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、
前記確率分布を平均化する段階と、

前記ノードのセットのランク付けに使用するため平均化した前記確率分布を伝える段階
とを有する、付記 1 に記載の方法。

(付記 9) 前記セッションノードから相異なる所定リンク数内のすべてのリンクされた
ノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

20

ユーザにより指定された嗜好に基づき前記ノードのセットのランク付けに使用する前記
複数の確率分布を伝える段階とをさらに有する、付記 1 に記載の方法。

(付記 10) 前記ノードのセットの前記確率分布を生成する段階は、一様な確率分布か
ら始めて確率分布を反復的に計算し、生成される前記確率分布は反復的に計算された確率
分布の収束の結果得られる段階を含む、付記 1 に記載の方法。

(付記 11) コンピュータにより実行される、有体的媒体にエンコードされた論理であ
って、前記論理は実行されると、

セッションノードと前記セッションノードにリンクされた複数のリンクされたノードと
を含むノードのセットのモデルにアクセスし、前記リンクされたノードは親ノードと子ノ
ードとを含み、親ノードは 1 つ以上の子ノードを前記セッションノードにリンクし、子ノ
ードは前記子ノードを前記セッションノードにリンクする 1 つ以上の親ノードを有する段
階と、

30

前記ノードのセットに対して、前記セッションノードから所定リンク数内のすべてのリ
ンクされたノードに確率を分配する確率分布を生成し、各子ノードはその親ノードの各々
から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノード
の各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する段階と、

前記ノードのセットのランク付けに使用するため前記確率分布を伝える段階とを実行す
る論理。

(付記 12) 前記ノードはハイパーテキスト・マークアップ・ランゲージ (HTML) 文
書である、付記 11 に記載の論理。

40

(付記 13) 前記所定数のリンクはユーザの我慢強さをモデル化したものである、付記
11 に記載の論理。

(付記 14) 前記所定数のリンクは 2 つのリンクである、付記 11 に記載の論理。

(付記 15) 前記所定割合は $1/2$ である、付記 11 に記載の論理。

(付記 16) 前記ノードのセットのモデルはワールド・ワイド・ウェブの少なくとも一部
をモデル化するグラフを含む、付記 11 に記載の論理。

(付記 17) 前記ランク付けは前記ノードのセットの検索を容易にして、ユーザからの
クエリに対する応答を生成する、付記 11 に記載の論理。

(付記 18) 実行されたとき、

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確

50

率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

前記確率分布を平均化する段階と、

前記ノードのセットのランク付けに使用するため平均化した前記確率分布を伝える段階とをさらに実行する、付記 11 に記載の論理。

(付記 19) 実行されたとき、

前記セッションノードから相異なる所定リンク数内のすべてのリンクされたノードに確率を分配する、前記ノードのセットの複数の確率分布を生成する段階と、

ユーザにより指定された嗜好に基づき前記ノードのセットのランク付けに使用する前記複数の確率分布を伝える段階とをさらに実行する、付記 11 に記載の論理。

(付記 20) 前記ノードのセットの前記確率分布を生成する段階は、一様な確率分布から始めて確率分布を反復的に計算し、生成される前記確率分布は反復的に計算された確率分布の収束の結果得られる段階を含む、付記 11 に記載の論理。

(付記 21) セッションノードと前記セッションノードにリンクされた複数のリンクされたノードとを含むノードのセットのモデルにアクセスし、前記リンクされたノードは親ノードと子ノードとを含み、親ノードは 1 つ以上の子ノードを前記セッションノードにリンクし、子ノードは前記子ノードを前記セッションノードにリンクする 1 つ以上の親ノードを有する手段と、

前記ノードのセットに対して、前記セッションノードから所定リンク数内のすべてのリンクされたノードに確率を分配する確率分布を生成し、各子ノードはその親ノードの各々から前記親ノードに分配された確率の所定割合を受け取り、前記親ノードはその子ノードの各々に前記親ノードに分配された前記確率の前記所定割合を一様に分配する手段と、

前記ノードのセットのランク付けに使用するため前記確率分布を伝える手段とを有するシステム。

【図面の簡単な説明】

【0021】

【図 1】ウェブページにわたるユーザの検索行動例を示す図である。

【図 2】ウェブページの第 1 のレイヤにわたる確率分布例を示す図である。

【図 3】図 2 に示したウェブページの第 2 のレイヤにわたる確率分布例を示す図である。

【図 4】図 2 に示したウェブページの第 1 のレイヤと第 2 のレイヤとにわたる確率分布例を示す図である。

【図 5】セッションベースクエリのためにノードのランク付けをするシステム例を示す図である。

【図 6】セッションベースクエリのためにノードのランク付けをする方法例を示す図である。

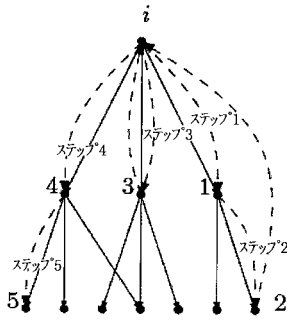
10

20

30

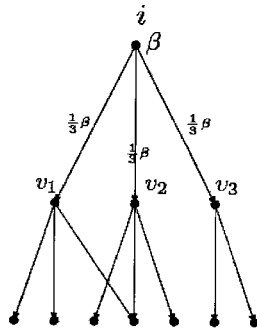
【図1】

ウェブページにわたるユーザの検索行動例を示す図



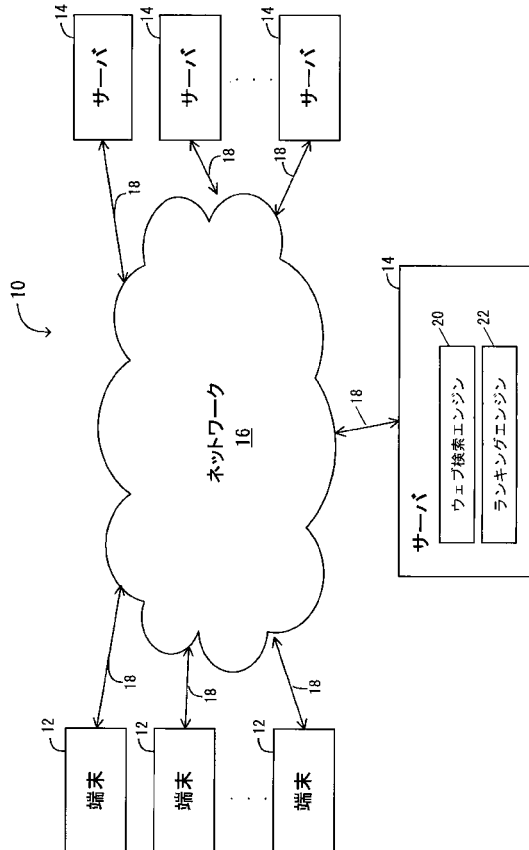
【図2】

ウェブページの第1のレイヤにわたる確率分布例を示す図



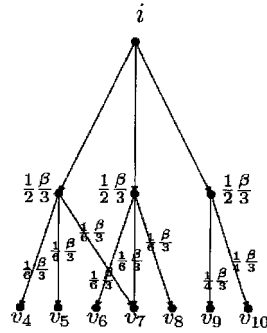
【図5】

セッションベーススクエリのためにノードのランク付けをするシステム例を示す図



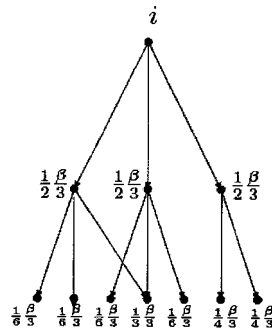
【図3】

図2に示したウェブページの第2のレイヤにわたる確率分布例を示す図



【図4】

図2に示したウェブページの第1のレイヤと第2のレイヤにわたる確率分布例を示す図



【図6】

セッションベーススクエリのためにノードのランク付けをする方法例を示す図

- 開始
- ↓
- 100 ノードのセットに対して一様な確率分布を生成する
- ↓
- 102 上位ノードから所定数のエッジ内で到達可能な下位ノードにその上位ノードの確率を分配して、確率分布を反復的に生成する
- ↓
- 104 最新の確率分布が前の確率分布に近いとき終了する
- ↓
- 106 ノードをランク付けするために、結果として得られた確率分布を伝える
- ↓
- 終了

フロントページの続き

審査官 伊知地 和之

(56)参考文献 米国特許第06285999(US, B1)

特開2006-107432(JP, A)

特開2006-127529(JP, A)

中窪 仁 外1名, “Web検索におけるリンク構造解析を利用したランキング法”, 情報処理学会研究報告, 日本, 社団法人情報処理学会, 2004年 7月15日, 第2004巻, 第72号, p. 411 - 415

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

CSD B(日本国特許庁)