

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成20年6月26日(2008.6.26)

【公開番号】特開2005-327293(P2005-327293A)

【公開日】平成17年11月24日(2005.11.24)

【年通号数】公開・登録公報2005-046

【出願番号】特願2005-141125(P2005-141125)

【国際特許分類】

G 0 6 F 17/30 (2006.01)

【 F I 】

G 0 6 F 17/30 1 8 0 Z

【手続補正書】

【提出日】平成20年5月13日(2008.5.13)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

複数のデータオブジェクトの属性の属性値を決定するコンピュータシステムにおいて実行される方法であって、

前記方法は、

前記複数のデータオブジェクトのタイプを規定するステップであって、各データオブジェクトのタイプがデータオブジェクトのタイプに固有の属性を有するステップと、

複数のデータオブジェクトを識別するステップであって、各データオブジェクトが一つのタイプに関連付けられるステップと、

前記複数のデータオブジェクトのタイプの各々に対して、

該データオブジェクトのタイプに関連付けられる複数のデータオブジェクト間の該データオブジェクトのタイプについての関係を識別し、

該データオブジェクトのタイプに関連付けられる一つのデータオブジェクトと他のデータオブジェクトのタイプに関連付けられる一つのデータオブジェクトとの間の該データオブジェクトのタイプについての関係を識別するステップと、

前記複数のデータオブジェクトのタイプの各々に対して、該データオブジェクトのタイプの属性の属性値を計算する各タイプについての関数を使用して前記識別された関係に基づいて前記データオブジェクトの属性の属性値を計算するステップであって、前記関数は、 F_i がデータオブジェクト i に関連付けられた属性値を表し、 R_i はデータオブジェクト i のタイプを持つ複数のデータオブジェクト間のタイプ内の関係を表し、および R_{j_i} はデータオブジェクト i のタイプを持つデータオブジェクトと他のデータオブジェクト j のタイプを持つデータオブジェクトとの間のタイプ間の関係を表す、

$F_i = F_i R_i + \sum_j R_{j_i} F_j R_{j_i}$

として定義されるステップと

前記計算された属性値を格納するステップと

を含み、

前記タイプ内の関係は、同じタイプのデータオブジェクト間の関係であり、

前記タイプ間の関係は、異なるタイプのデータオブジェクト間の関係であることを特徴とする方法。

【請求項2】

複数のデータオブジェクトの属性の属性値を決定するコンピュータシステムにおいて実行される方法であって、

前記方法は、

前記複数のデータオブジェクトのタイプを規定するステップであって、各データオブジェクトのタイプがデータオブジェクトのタイプに固有の属性を有するステップと、

複数のデータオブジェクトを識別するステップであって、各データオブジェクトが一つのタイプに関連付けられるステップと、

前記複数のデータオブジェクトのタイプの各々に対して、

該データオブジェクトのタイプに関連付けられる複数のデータオブジェクト間の該データオブジェクトのタイプについての関係を識別し、

該データオブジェクトのタイプに関連付けられる一つのデータオブジェクトと他のデータオブジェクトのタイプに関連付けられる一つのデータオブジェクトとの間の該データオブジェクトのタイプについての関係を識別するステップと、

前記複数のデータオブジェクトのタイプの各々に対して、該データオブジェクトのタイプの属性の属性値を計算する各タイプについての関数を使用して前記識別された関係に基づいて前記データオブジェクトの属性の属性値を計算するステップであって、前記関数は、

【数 1】

$$w_M = \alpha_M L_M^T w_M + \beta_{NM} \sum_{\forall N=M} L_{NM}^T w_N$$

であるステップと、

前記計算された属性値を格納するステップと

を含み、

前記タイプ内の関係は、同じタイプのデータオブジェクト間の関係であり、

前記タイプ間の関係は、異なるタイプのデータオブジェクト間の関係であることを特徴とする方法。

【請求項 3】

前記データオブジェクトのタイプは、権威タイプ、ハブタイプ及び専門知識タイプを含むことを特徴とする請求項 1 または請求項 2 に記載の方法。

【請求項 4】

前記権威タイプの複数のデータオブジェクト間の関係は、Web ページが別の Web ページへのリンクを有するかどうかに基づいて定まることを特徴とする請求項 3 に記載の方法。

【請求項 5】

前記権威タイプのデータオブジェクトと前記専門知識タイプのデータオブジェクトとの間の関係は、ユーザによる Web ページへのアクセスに基づいて定まることを特徴とする請求項 4 に記載の方法。

【請求項 6】

前記ハブタイプの複数のデータオブジェクト間の関係は、Web ページが別の Web ページへのリンクを有するかどうかに基づいて定まることを特徴とする請求項 3 に記載の方法。

【請求項 7】

前記ハブタイプのデータオブジェクトと前記専門知識タイプのデータオブジェクトとの間の関係は、ユーザによる Web ページへのアクセスに基づいて定まることを特徴とする請求項 6 に記載の方法。

【請求項 8】

あるタイプのデータオブジェクトを該オブジェクトの属性値に基づいて格付けするステップをさらに含むことを特徴とする請求項 1 または請求項 2 に記載の方法。

【請求項 9】

前記データオブジェクトのタイプの属性値を定義する各データオブジェクトのタイプについての式を規定するステップをさらに含むことを特徴とする請求項 1 または請求項 2 に記載の方法。

【請求項 10】

前記計算するステップは、前記式を繰り返し解くステップを含むことを特徴とする請求項 9 に記載の方法。

【請求項 11】

前記式は、他の式の属性値に基づいて再帰的に定義されることを特徴とする請求項 9 に記載の方法。

【請求項 12】

複数のデータオブジェクトの属性値を決定するコンピュータシステムにおいて実行される方法であって、

前記方法は、

あるタイプのデータオブジェクトについてのタイプに固有の属性の属性値を、該タイプのデータオブジェクトと、別のタイプに固有の属性を有する別のタイプのデータオブジェクトとの間の関係に基づいて計算する関数を規定するステップであって、前記規定された関数は、該タイプの複数のデータオブジェクト間の関係にも基づいて計算され、前記関数は、 F_i がデータオブジェクト i に関連付けられた属性値を表し、 R_i はデータオブジェクト i のタイプを持つ複数のデータオブジェクト間のタイプ内の関係を表し、および R_{j_i} はデータオブジェクト i のタイプを持つデータオブジェクトと他のデータオブジェクト j のタイプを持つデータオブジェクトとの間のタイプ間の関係を表す、

$F_i = F_i R_i + \sum_j R_{j_i} F_j R_{j_i}$
として定義されるステップと、

該タイプのデータオブジェクトと前記別のタイプのデータオブジェクトとの間の関係を特定するデータを受け取るステップと、

前記規定された関数を計算して、該タイプの複数のデータオブジェクトの属性値を決定するステップと、

前記決定された属性値を格納するステップと

を含み、

前記タイプ内の関係は、同じタイプのデータオブジェクト間の関係であり、

前記タイプ間の関係は、異なるタイプのデータオブジェクト間の関係であることを特徴とする方法。

【請求項 13】

前記別のタイプのデータオブジェクトのタイプに固有の属性の属性値を計算する関数を規定するステップをさらに含むことを特徴とする請求項 12 に記載の方法。

【請求項 14】

前記関数は、再帰的に定義されることを特徴とする請求項 13 に記載の方法。

【請求項 15】

前記計算するステップは、前記属性値が解に収束するまで各関数を繰り返し計算するステップを含むことを特徴とする請求項 14 に記載の方法。

【請求項 16】

前記関数は、一次方程式を表すことを特徴とする請求項 14 に記載の方法。

【請求項 17】

コンピュータ読み取り可能な記録媒体であって、

前記記録媒体は、

第 1 のタイプのデータオブジェクトに対する第 1 のタイプに固有の属性の属性値を、前記第 1 のタイプのデータオブジェクトと、第 2 のタイプに固有の属性を有する第 2 のタイプのデータオブジェクトとの間の関係に基づいて計算する第 1 の関数を規定するステップと、

前記第 2 のタイプのデータオブジェクトに対する前記第 2 のタイプに固有の属性の属性

値を、前記第 2 のタイプの複数のデータオブジェクト間の関係に基づいて計算する第 2 の関数を規定するステップと、

前記第 1 のタイプのデータオブジェクトと前記第 2 のタイプのデータオブジェクトとの間の関係を特定するデータを受け取るステップと、

前記第 1 のタイプのデータオブジェクトの属性値および前記第 2 のタイプのデータオブジェクトの属性値を決定する前記規定された第 1 の関数および第 2 の関数を計算するステップであって、前記関数は、 F_i がデータオブジェクト i に関連付けられた属性値を表し、 R_i はデータオブジェクト i のタイプを持つ複数のデータオブジェクト間のタイプ内の関係を表し、および R_{j_i} はデータオブジェクト i のタイプを持つデータオブジェクトと他のデータオブジェクト j のタイプを持つデータオブジェクトとの間のタイプ間の関係を表す、

$F_i = F_i R_i + \sum_j R_{j_i} F_j R_{j_i}$
として定義されるステップと

前記決定された属性値を格納するステップと

を含む方法によってデータオブジェクトの属性値を決定するコンピュータシステムを制御する命令を含み、

前記タイプ内の関係は、同じタイプのデータオブジェクト間の関係であり、

前記タイプ間の関係は、異なるタイプのデータオブジェクト間の関係であることを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項 18】

前記第 1 の関数は、前記第 1 のタイプの複数のデータオブジェクト間の関係に基づいて属性値をさらに計算することを特徴とする請求項 17 に記載のコンピュータ読み取り可能な記録媒体。

【請求項 19】

前記第 2 の関数は、前記第 1 のタイプのデータオブジェクトの属性値に基づいて属性値を計算することを特徴とする請求項 17 に記載のコンピュータ読み取り可能な記録媒体。

【請求項 20】

前記関数は、再帰的に定義されることを特徴とする請求項 17 に記載のコンピュータ読み取り可能な記録媒体。

【請求項 21】

前記計算するステップは、前記属性値が解に収束するまで各関数を繰り返し計算するステップを含むことを特徴とする請求項 20 に記載のコンピュータ読み取り可能な記録媒体。

【請求項 22】

前記関数は、一次方程式を表すことを特徴とする請求項 17 に記載のコンピュータ読み取り可能な記録媒体。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【発明の名称】タイプ内およびタイプ間の関係に基づいてオブジェクトを格付けする方法およびシステム

【技術分野】

【0001】

本発明は、一般には、オブジェクトの格付けに関し、より詳細には、オブジェクト関係に基づく格付けに関する。

【背景技術】

【0002】

GoogleおよびOvertureなど多くの検索エンジンサービスは、インターネットを介してアクセス可能な情報の検索を提供する。これらの検索エンジンサービスによりユーザは、Webページなど、ユーザにとって興味深いと思われる表示ページを検索することができる。検索語を含む検索要求（「クエリ」とも呼ばれる）をユーザが送信した後、検索エンジンサービスは、これらの検索語に関連していると思われるWebページを識別する。関連するWebページを迅速に識別するために、検索エンジンサービスは、Webページへのキーワードのマッピングを保持することができる。検索エンジンサービスは、Web（すなわちワールドワイドウェブ）を「クロール」することによりこのマッピングを生成して、各Webページのキーワードを抽出することができる。Webをクロールするために、検索エンジンサービスは、ルートのWebページのリストを使用し、これらのルートのWebページを通じてアクセス可能なすべてのWebページを識別することができる。特定のWebページのキーワードは、見出し語、Webページのメタデータで供給された語、ハイライト表示された語などを識別するような、様々な周知の情報検索技術を使用して抽出することができる。検索エンジンサービスにより、各々の合致の近さ、Webページの人気（例えばGoogleのPageRank）などに基づいて各Webページが検索要求にどの程度関連しているかを示す適合性の評点（score）を計算することができる。次に、検索エンジンサービスにより、これらのWebページへのリンクを適合性に基づいた順序でユーザに表示する。検索エンジンにより、さらに一般的に、あらゆる文書の集まりで情報の検索を提供することができる。例えば、文書の集まりには、すべての米国特許、すべての連邦の法廷意見、企業のすべての保管文書などを含めることもできる。

【0003】

Webページを格付けするための2つの周知の技術は、PageRankおよびHITS（「Hyperlinked-Induced Topic Search」）である。PageRankは、Webページが重要なWebページへのリンク（つまり「外向きのリンク」）を持つという原則に基づく。したがって、Webページの重要度は、そのWebページにリンク（つまり「内向きのリンク」）する他のWebページの数および重要度に基づく。簡単な形式では、Webページ間のリンクは行列Aによって表すことができ、ここで A_{ij} はWebページiからWebページjへの外向きのリンクの数を表す。Webページjの重要度の評点 w_j は、以下の式で表すことができる。

$$w_j = \sum_i A_{ij} w_i$$

この式は、以下の式に基づく反復計算によって解くことができる。

$$A^T w = w$$

ただし、 w はWebページの重要度の評点のベクトルであり、 A^T の優固有ベクトル（principal eigenvector）である。

【0004】

HITS技術はさらに、他の重要なWebページへのリンクを数多く持つWebページはそれ自体が重要であるという原則に基づく。したがって、HITSは、Webページの「重要度」を、「ハブ」および「権威（authority）」という2つの関連する属性に分割する。「ハブ」は、WebページがリンクするWebページの「権威」の評点によって測定され、「権威」は、WebページにリンクするWebページの「ハブ」の評点によって測定される。Webページの重要度をクエリからは独立して計算するPageRankとは対照的に、HITSでは、内向きおよび外向きのリンクをたどることによって結果のWebページおよび結果のWebページに関連するWebページに基づいて重要度を計算する。HITSは、検索エンジンサービスにクエリを送信し、結果のWebページをWebページの初期セットとして使用する。HITSは、内向きのリンクの宛先であるWebページおよび結果のWebページの外向きのリンクのソースであるWebページをセットに追加する。次にHITSは、反復アルゴリズムを使用して各Webページの権威およびハブの評点を計算する。権威およびハブの評点は、以下の式で表すことができる。

【0005】

【数 2】

$$a(p) = \sum_{q \rightarrow p} h(q) \quad \text{および} \quad h(p) = \sum_{p \rightarrow q} a(q)$$

【0006】

ただし、 $a(p)$ は Web ページ p の権威の評点を表し、 $h(p)$ は Web ページ p のハブの評点を表す。HITS では、隣接行列 A を使用してリンクを表す。隣接行列は、以下の式で表すことができる。

【0007】

【数 3】

$$b_{ij} = \begin{cases} 1 & \text{ページ } i \text{ がページ } j \text{ へのリンクを有する場合、} \\ 0 & \text{ページ } i \text{ がページ } j \text{ へのリンクを有さない場合} \end{cases}$$

【0008】

ベクトル a および h はそれぞれ、セット内のすべての Web ページの権威およびハブの評点に対応し、以下の式で表すことができる。

$$a = A^T h \quad \text{および} \quad h = A a$$

したがって、 a および h は、行列 $A^T A$ および $A A^T$ の固有ベクトルである。HITS はまた、アクセス数によって測定される Web ページの人気を計算に入れるように変更することもできる。Web ログの分析に基づいて、隣接行列の b_{ij} は、ユーザが Web ページ i から Web ページ j に移動するときに増分することができる。

【発明の開示】

【発明が解決しようとする課題】

【0009】

上述の Web ページ格付け技術では、Web ページの格付けを主として Web ページ自体の属性に基づいて行う。Web ページ自体の属性は、一方の Web ページから別の Web ページへのリンクを含み、一方の Web ページから別の Web ページへ伝わる。格付け技術では、Web ページに直接関連がない属性を計算に入れることができない。例えば、Web ページの重要度は、Web ページにアクセスするユーザの専門知識が計算に入れられる場合、より正確に決定することができる。Web ページに直接関連がない属性に基づいて Web ページの重要度を計算するための技術を有することが望ましい。より一般的には、一方のタイプのオブジェクト（例えば、Web ページ）の評点を、別のタイプのオブジェクト（例えば、ユーザ）との関係に基づいて生成することが望ましい。

【課題を解決するための手段】

【0010】

異なるオブジェクトタイプのオブジェクトとの関係に基づいてオブジェクトを格付けするための方法およびシステムを提供する。格付けシステムにより、オブジェクトの各タイプの属性ごとに式を定義する。式は、属性値を定義し、属性と、同じタイプのオブジェクトおよび異なるタイプのオブジェクトに関連付けられる属性との間の関係に基づく。属性値は、一方の属性が別の属性に関して定義され、逆の場合も同じであるように相互依存することができるので、式は属性の再帰的定義を表す。格付けシステムにより、属性値が解に収束するまで式を使用して、オブジェクトの属性値を繰り返し計算する。次に、格付けシステムにより、属性値に基づいてオブジェクトの格付けを行う。

【発明を実施するための最良の形態】

【0011】

同じまたは別のデータオブジェクトタイプのデータオブジェクトとの関係に基づいてデータオブジェクトタイプのデータオブジェクトを格付けするための方法およびシステムを提供する。一実施形態において、格付けシステムにより、様々なデータオブジェクトタイプのデータオブジェクトを識別する。例えば、一方のデータオブジェクトタイプが Web ページであり、別のデータオブジェクトタイプがクエリであり、および別のデータオブジェクトタイプがユーザである場合がある。各データオブジェクトタイプは、様々なタイプ

固有の属性を有することができる。例えば、Web ページが権威属性を有し、およびユーザが専門知識属性を有する場合がある。Web ページの権威属性は、Web ページの内向きのリンクの数に基づくことができる。ユーザの専門知識属性は、ユーザが高い権威属性値を持つ Web ページにアクセスする場合に増加される。格付けシステムにより、データオブジェクトの属性値を計算し、およびデータオブジェクトの属性値に基づいてデータオブジェクトを格付けすることができる。

【0012】

格付けシステムにより、各オブジェクトが単一の属性を含むように、オブジェクトの「タイプ」を定義する。例えば、格付けシステムにより、Web ページの権威属性に対応する タイプ、および Web ページのハブ属性に対応する別の タイプ を定義することができる。したがって、2つの タイプ は、同じ基礎をなすデータオブジェクト（例えば Web ページ）を表すことができる。格付けシステムにより、タイプ 内の関係と呼ばれる同じ タイプ のオブジェクトの間の様々な関係、および タイプ 間の関係と呼ばれる異なる タイプ のオブジェクトの間の様々な関係を決定する。例えば、クエリが送信される場合、格付けシステムにより、結果を権威 タイプ のオブジェクトとして使用し、および Web ログを使用して、これらの Web ページにアクセスしたユーザを専門知識 タイプ のオブジェクトとして識別することができる。権威 タイプ の タイプ 内の関係のオブジェクトは、Web ページの内向きのリンクと外向きのリンクとの関係を含むことができる。例えば、Web ページが別の Web ページへのリンクを有する場合、Web ページはその他の Web ページへの外向きのリンクの関係を有し、およびその他の Web ページは Web ページへの内向きのリンクの関係を有する。権威 タイプ と専門知識 タイプ とのオブジェクトの間の タイプ 間の関係は、Web ページへのユーザアクセスに基づく。例えば、ユーザが Web ページにアクセスする場合、Web ページおよびユーザはアクセスの関係を有する。格付けシステムにより、他の タイプ のオブジェクトの属性値と組み合わせられた タイプ 内の関係および タイプ 間の関係を使用して、所定の タイプ のオブジェクトに対する属性の値を導き出す。例えば、格付けシステムにより、内向きと外向きとのリンクの関係およびユーザアクセスの関係をを使用して、Web ページの権威およびハブ属性ならびにユーザの専門知識属性を導き出すことができる。

【0013】

一実施形態において、格付けシステムにより、一次方程式など一連の式を使用して関係および属性を表す。格付けシステムにより、別の タイプ の属性値に基づいて再帰的に定義される一次方程式を使用して、各 タイプ の属性値を表す。例えば、権威属性の一次方程式が専門知識属性の属性値に基づいて定義されることも、またその逆の場合もある。一次方程式は再帰的に定義することができるので、格付けシステムにより、属性値が解に収束するまで、各一次方程式の属性値を繰り返し計算することによって一次方程式を解く。一次方程式を解いた後、格付けシステムにより、属性値に基づいてデータオブジェクトを格付けすることができる。例えば、格付けシステムにより、Web ページの権威属性値に基づいて Web ページを格付けすることができる。

【0014】

格付けシステムにより、オブジェクトの タイプ 内と タイプ 間との関係に基づいて、オブジェクトの属性値を表す。属性の値は、以下の数式によって表すことができる。

$$F_i = F_i R_{i j} + F_j R_{j i}$$

【0015】

ただし、 F_i はオブジェクト i に関連付けられる属性値を表し、 $R_{i j}$ はオブジェクト i の タイプ のオブジェクトの間の タイプ 内関係を表し、 $R_{j i}$ はオブジェクト i の タイプ と他のオブジェクト j の タイプ とのオブジェクトの間の タイプ 間の関係を表す。 $x = \{x_1, x_2, \dots, x_m\}$ および $y = \{y_1, y_2, \dots, y_n\}$ の2つのオブジェクトの タイプ がある場合、これらの タイプ 内の関係は R_x および R_y として表すことができ、タイプ 間の関係は $R_{x y}$ および $R_{y x}$ によって表すことができる。格付けシステムでは、隣接行列を使用して関係情報を表す。 L_x および L_y は、それぞれ集合 X および Y の タイプ 内の

関係の隣接行列を表す。 $L_{X,Y}$ および $L_{Y,X}$ は、それぞれ、 X のオブジェクトから Y のオブジェクトへの タイプ間の関係 の隣接行列および Y のオブジェクトから X のオブジェクトへの タイプ間の関係 の隣接行列を表す。格付けシステムでは、以下のように隣接行列を表す。

【 0 0 1 6 】

【 数 4 】

$$L_{XY}(i, j) = \begin{cases} 1 & \text{オブジェクト } x_i \text{ からオブジェクト } y_j \text{ へ関係がある場合} \\ 0 & \text{オブジェクト } x_i \text{ からオブジェクト } y_j \text{ へ関係がない場合} \end{cases}$$

【 0 0 1 7 】

ただし、 $L_{X,Y}(i, j)$ は、集合 X のオブジェクト i から集合 Y のオブジェクト j への関係（「リンク」とも呼ばれる）が存在するかどうかを示す。属性値の一次方程式は、以下の式のように表すことができる。

【 0 0 1 8 】

【 数 5 】

$$W_Y = L_Y^T W_Y + L_{XY}^T W_X \quad (1)$$

$$W_X = L_X^T W_X + L_{YX}^T W_Y$$

【 0 0 1 9 】

ただし、 w_X は、 X のオブジェクトの属性ベクトルであり、 w_Y は Y のオブジェクトの属性ベクトルである。式 1 は、以下の形式に一般化することができる。

【 0 0 2 0 】

【 数 6 】

$$W_M = L_M^T W_M + \sum_{\forall N \neq M} L_{NM}^T W_N \quad (2)$$

【 0 0 2 1 】

ただし、 M は属性ベクトルの行列を表す。

【 0 0 2 2 】

オブジェクトの間の相互に補強し合う関係がオブジェクトに過度の属性値をもたらすことがあるので、格付けシステムにより、オブジェクトが 1 つの隣接行列内の n 個の他のオブジェクトに関連する場合、各関連するオブジェクトはその $1/n$ 番目の属性値を受け取るような方法で、バイナリの隣接行列を正規化することができる。格付けシステムにより、Page Rank のランダムサーファーマデルを導入して、ランダムな関係をシミュレートし、したがって、以下に説明する計算時のシンクノードを回避することができる。さらに、異なる タイプ からの属性が各々他の属性に対して異なる重要度を有する可能性があるため、格付けシステムにより、タイプ の各々の組み合わせに重み付けを使用することができる。したがって、格付けシステムにより、正規化、ランダムサーファーマデル、および重み付けを計算に入れて、以下の式によって属性値を表すことができる。

【 0 0 2 3 】

【 数 7 】

$$W_M = \alpha_M L_M'^T W_M + \beta_{NM} \sum_{\forall N \neq M} L_{NM}'^T W_N$$

ただし、

$$\alpha_M + \sum_{\forall N \neq M} \beta_{NM} = 1; \alpha_M > 0 \beta_{NM} > 0; \quad (3)$$

$$L_M' = \varepsilon U + (1 - \varepsilon) L_M; 0 < \varepsilon < 1;$$

$$L_{NM}' = \delta_N U + (1 - \delta_N) L_{NM}; 0 < \delta_N < 1;$$

【 0 0 2 4 】

ただし、 U は一様な推移確率の推移行列であり（すべての i, j について $U_{ij} = 1/n$ ）

、ただし、 n はデータ空間 N のオブジェクトの合計数)、 L_M および L_{NM} は正規化された隣接行列であり、 α および β は行列 L_M および L_{NM} 内のランダムな関係をシミュレートするために使用される平滑化因数であり、 α_M および β_{NM} は関係の重み付けを表す。格付けシステムにより、収束するまで式3を繰り返し計算する。式3は、以下の式で表される統合された正方行列 A によって表すことができる。

【0025】

【数8】

$$A = \begin{pmatrix} \alpha_1 L'_1 & \beta_{12} L'_{12} & \cdots & \beta_{1n} L'_{1n} \\ \beta_{21} L'_{21} & \alpha_2 L'_2 & \cdots & \beta_{2n} L'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} L'_{m1} & \beta_{n2} L'_{n2} & \cdots & \alpha_n L'_n \end{pmatrix} \quad (4)$$

【0026】

行列 A は、対角線上に

【0027】

【数9】

L'_M

【0028】

、統合された行列の他の部分に

【0029】

【数10】

L'_{NM}

【0030】

を有する。格付けシステムにより、行列 A を使用して、異なるデータ空間にあるすべてのデータオブジェクトの属性ベクトルである、ベクトル w を変換するために反復アプローチを使用する(例えば、 $w = A^T w$)。反復が収束する場合、ベクトル w は行列 A の優固有ベクトルである。

【0031】

M および N が異質のデータ空間である場合、格付けシステムでは、ランダムな関係を使用して無関係を表す。 M のオブジェクトが N のオブジェクトにリンクの関係を持たない場合、部分行列

【0032】

【数11】

L'^T_{NM}

【0033】

はゼロであり、「シンクノード」を表す。計算では、シンクノードにすべての属性値を割り当てることがある。これを防ぐために、格付けシステムにより、部分行列

【0034】

【数12】

L'^T_{NM}

【0035】

の対応する行のすべての要素を $1/n$ に設定する。ただし、 n はデータ空間 N のオブジェクトの合計数である。代替として、格付けシステムにより、望ましくないタイプ内およびタイプ間の関係に対して、対応する重み付けを0に設定することができる。しかし、 α_M が0よりも大きい場合、反復計算が収束するのであれば β_{NM} は0よりも大きい必要がある。したがって、

【 0 0 3 6 】

【 数 1 3 】

$L'_{NM}{}^T$

【 0 0 3 7 】

の関係が望ましくない場合、格付けシステムにより、 N_M を非常に小さい正の重みに設定して

【 0 0 3 8 】

【 数 1 4 】

$L'_{NM}{}^T$

【 0 0 3 9 】

の影響を低減する。

【 0 0 4 0 】

すべての隣接行列を使用して統合された行列を構築することにより、格付けシステムでは、異なるオブジェクトのタイプを含む、統合されたデータ空間を構築する。したがって、以前のタイプ間の関係は、統合された空間のタイプ間の関係と見なすことができ、格付けシステムにより、効率的に単一のデータ空間のリンク解析を行う。

【 0 0 4 1 】

図 1 は、一実施形態における格付けシステムのコンポーネントを例示するフロー図である。格付けシステム 110 は、通信リンク 102 を介して様々な Web サイト 101 に接続されている。格付けシステムには、オブジェクト収集コンポーネント 112、関係確立コンポーネント 113、評点計算コンポーネント 114、およびオブジェクト順序付けコンポーネント 115 を呼び出してオブジェクトを格付けするオブジェクト格付けコンポーネント 111 を含む。オブジェクト格付けコンポーネントは、Web ページのセットを受け取り、タイプ内およびタイプ間の関係に基づいて Web ページを格付けすることができる。オブジェクト収集コンポーネントは、様々なタイプのオブジェクトに関連する関係情報を取り出す。例えば、オブジェクト収集コンポーネントは、Web サイトの Web ログにアクセスして、どのユーザがどの Web ページにアクセスするのかを識別することができる。関係確立コンポーネントは、タイプ内およびタイプ間の関係の行列を作成する。例えば、関係の行列は、ユーザがアクセスする Web ページにユーザをマッピングすることができる。評点計算コンポーネントは、属性値が解に収束するまで、式 3 を使用して属性値を再帰的に計算する。オブジェクト順序付けコンポーネントは、属性値に基づいてデータオブジェクトをソートする。例えば、オブジェクト順序付けコンポーネントは、Web ページの権威属性の値を使用して、Web ページをソートすることができる。

【 0 0 4 2 】

格付けシステムが実装されるコンピューティング装置には、中央処理装置、メモリ、入力装置（例えば、キーボードおよびポインティングデバイスなど）、出力装置（例えば、ディスプレイ装置など）、および記憶装置（例えば、ディスクドライブなど）を含めることができる。メモリおよび記憶装置は、格付けシステムを実装する命令を含めることができるコンピュータ読取可能な媒体である。さらに、データ構造およびメッセージ構造を、通信リンク上の信号などのデータ伝送媒体を介して格納または送信することができる。インターネット、ローカルエリアネットワーク、ワイドエリアネットワーク、またはポイントツーポイントのダイヤルアップ接続など、様々な通信リンクを使用することができる。

【 0 0 4 3 】

格付けシステムを、様々なオペレーティング環境において実装することができる。使用に最適な様々な周知のコンピューティングシステム、環境、および構成には、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルドまたはラップトップ装置、マルチプロセッサシステム、マイクロプロセッサに基づくシステム、プログラム可能な家庭用電化製品、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータ、上記のシステ

ムまたは装置のいずれかを含む分散コンピューティング環境などが含まれる。

【 0 0 4 4 】

格付けシステムにより、1つまたは複数のコンピュータまたは他の装置によって実行されるプログラムモジュールなど、コンピュータ実行可能命令の一般的なコンテキストにおいて説明することができる。一般に、プログラムモジュールには、特定のタスクを実行するかまたは特定の抽象データ型を実装するルーチン、プログラム、オブジェクト、コンポーネント、データ構造などが含まれる。通常、プログラムモジュールの機能を、様々な実施形態において望ましいように組み合わせまたは分散することができる。

【 0 0 4 5 】

図2は、一実施形態におけるオブジェクト格付けコンポーネントの処理を例示するフロー図である。コンポーネントは、オブジェクト情報を収集し、オブジェクトの間の関係を確立し、オブジェクトの属性値を計算し、属性に基づいてオブジェクトを順序付ける。ブロック201において、コンポーネントは、様々なオブジェクトに関連する情報を収集する。ブロック202において、コンポーネントは、関係確立コンポーネントを呼び出して、隣接行列を生成する。関係確立コンポーネントはまた、およびの重みを取り出して調整することもできる。ブロック203において、コンポーネントは、評点計算コンポーネントを呼び出して、属性値が解に収束するまで属性値を繰り返し計算する。ブロック204において、コンポーネントは、属性の値に基づいてデータオブジェクトを順序付ける。例えば、コンポーネントは、権威属性に基づいてWebページを順序付けることができる。

【 0 0 4 6 】

図3は、一実施形態における関係確立コンポーネントの処理を例示するフロー図である。ブロック301から303において、コンポーネントは、各タイプに対する隣接行列を確立するループを行う。ブロック301において、コンポーネントは次のタイプを選択する。判断ブロック302において、すべてのタイプがすでに選択されている場合、コンポーネントは戻り、別の場合、コンポーネントはブロック303において続行する。ブロック303において、コンポーネントは、選択されたタイプのオブジェクトとすべてのタイプのオブジェクトとの間の関係を確立する。例えば、コンポーネントは、権威タイプとハブタイプとの間の関係、および権威タイプと専門知識タイプとの間の関係を確立する。次に、コンポーネントはブロック301へのループを行い、次のタイプを選択する。

【 0 0 4 7 】

図4は、一実施形態における評点計算コンポーネントの処理を例示するフロー図である。コンポーネントは、属性値が収束するまで式を繰り返し計算する。ブロック401において、コンポーネントは、隣接行列によって表されたオブジェクトの関係を取り出す。ブロック402において、コンポーネントは、タイプ内およびタイプ間の関係に対する重みおよびを取り出す。ブロック403において、コンポーネントは、各タイプのベクトル w を初期化して、そのタイプの各オブジェクトに対し等しい属性値を持つようにする。コンポーネントは、各値を $1/m$ に設定することができる。ただし、 m はタイプのオブジェクトの数である。例えば、10のユーザがある場合、コンポーネントは専門知識タイプの初期の属性値を $1/10$ に設定する。コンポーネントはまた、判断ブロック405の検査を最初に通過するように、各タイプの差異変数を大きい値に初期化する。コンポーネントは、計算が解に収束しているかどうかを決定するために、各反復の最後において各差異変数に新しい値を計算する。ブロック404から409において、コンポーネントは、計算が解に収束するまで式3の計算を実行する。ブロック404において、コンポーネントは次の反復を開始する。判断ブロック405において、前回の反復中に計算された差異の合計が差異のしきい値よりも小さい場合、計算は解に収束しており、コンポーネントは戻り、別の場合、コンポーネントはブロック406において続行する。ブロック406において、コンポーネントは次のタイプを選択する。判断ブロック407において、すべてのタイプがすでに選択されている場合、コンポーネントはブロック404にループを行い、次の反復を開始し、別の場合、コンポーネントはブロック408において続行する。プロ

ック408において、コンポーネントは、以前の反復において計算された値に基づいて、選択されたタイプの値を計算する。ブロック409において、コンポーネントは、この反復の値と選択されたタイプの以前の反復の値との間の差異を計算する。次に、コンポーネントはブロック406にループを行い、次のタイプを選択する。

【0048】

本明細書において格付けシステムの特定の実施形態について例示のために説明しているが、本発明の精神および範囲を逸脱することなく、様々な変更を行うことができることを、当事業者であれば理解するであろう。例えば、当事業者であれば、属性値を表すために非線形方程式を使用することができることを理解するであろう。また、格付けシステムは、相互に何らかの関係を有するすべてのタイプのオブジェクトに使用することができる。例えば、格付けシステムにより、学生または志願者と教授との関係を使用する「重要度」に基づいて大学を格付けするために使用することもできる。ここで、大学、学生、および教授は異なるタイプのオブジェクトを表す。したがって、本発明は、添付の請求の範囲による場合を除き、限定されることはない。

【図面の簡単な説明】

【0049】

- 【図1】一実施形態における格付けシステムのコンポーネントを例示するフロー図である。
- 【図2】一実施形態におけるオブジェクト格付けコンポーネントの処理を例示するフロー図である。
- 【図3】一実施形態における関係確立コンポーネントの処理を例示するフロー図である。
- 【図4】一実施形態における評点計算コンポーネントの処理を例示するフロー図である。