

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5681801号
(P5681801)

(45) 発行日 平成27年3月11日(2015.3.11)

(24) 登録日 平成27年1月16日(2015.1.16)

(51) Int.Cl.

G06F 17/30 (2006.01)

F 1

G06F 17/30 210D
G06F 17/30 419A

請求項の数 14 (全 21 頁)

(21) 出願番号	特願2013-529133 (P2013-529133)
(86) (22) 出願日	平成23年9月20日 (2011.9.20)
(65) 公表番号	特表2013-541090 (P2013-541090A)
(43) 公表日	平成25年11月7日 (2013.11.7)
(86) 國際出願番号	PCT/US2011/001623
(87) 國際公開番号	W02012/039760
(87) 國際公開日	平成24年3月29日 (2012.3.29)
審査請求日	平成25年7月31日 (2013.7.31)
(31) 優先権主張番号	201010292384.0
(32) 優先日	平成22年9月21日 (2010.9.21)
(33) 優先権主張国	中国(CN)
(31) 優先権主張番号	13/200, 124
(32) 優先日	平成23年9月19日 (2011.9.19)
(33) 優先権主張国	米国(US)

(73) 特許権者	510330264 アリババ・グループ・ホールディング・リミテッド ALIBABA GROUP HOLDING LIMITED 英國領, ケイマン諸島, グランド・ケイマン, ジョージ・タウン, ワン・キャピタル・プレイス, フォース・フロア, ピー.オ.一. ボックス 847
(74) 代理人	110000028 特許業務法人明成国際特許事務所
(74) 代理人	100102989 弁理士 井上 佳知

最終頁に続く

(54) 【発明の名称】 カテゴリ分けされた製品情報の処理

(57) 【特許請求の範囲】

【請求項 1】

データ処理システムであって、

1つ以上のプロセッサであって、

複数の階層レベルを含む製品カテゴリツリーにアクセスするように構成され、

前記複数の階層レベルのリーフカテゴリレベルは、リーフカテゴリノードを含み、

前記リーフカテゴリノードは、製品情報を含み、

前記製品情報は、複数の製品属性パラメータと、複数の非製品属性パラメータと、を含み、

前記複数の製品属性パラメータおよび前記複数の非製品属性パラメータは、標準製品情報を構成し、

前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択するように構成され、

それぞれ前記標準製品情報の少なくとも一部を含む複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割するように構成され、

前記複数の非製品属性パラメータの類似度を決定するように構成され、

前記複数のセットのうちの1つのセット内の前記標準製品情報を、各サブセット内の前記類似度が閾値を満たすように複数のサブセットに分割するように構成されている1つ以上のプロセッサと、

10

20

前記 1 つ以上のプロセッサに接続され、前記 1 つ以上のプロセッサに命令を提供するよう
に構成されている 1 つ以上のメモリと、
を備える、システム。

【請求項 2】

請求項 1 に記載のシステムであって、
前記代表的な製品属性パラメータは、製品を検索するためにユーザによって以前に使用
されている、システム。

【請求項 3】

請求項 1 に記載のシステムであって、
前記複数のセットのうちの 1 つのセットは、前記セットに関連付けられている全ての製
品において同一の値の製品属性パラメータを含む、システム。 10

【請求項 4】

請求項 1 に記載のシステムであって、
前記プロセッサは、更に、製品のお勧めをするために前記複数のセットを使用するよう
に構成されている、システム。

【請求項 5】

請求項 1 に記載のシステムであって、
前記プロセッサは、更に、検索結果を提供するために前記複数のセットを検索するよう
に構成されている、システム。

【請求項 6】

請求項 1 に記載のシステムであって、
前記プロセッサは、更に、製品のお勧めをするために前記複数のサブセットを使用する
ように構成されている、システム。 20

【請求項 7】

請求項 1 に記載のシステムであって、
前記複数の製品属性パラメータは、更に、非標準製品情報を構成する、システム。

【請求項 8】

請求項 1 に記載のシステムであって、
前記複数の製品属性パラメータは、更に、非標準製品情報を構成し、
前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された製品
属性パラメータの値が設定基準を満たすかどうかを決定するように構成されている、シス
テム。 30

【請求項 9】

請求項 8 に記載のシステムであって、
前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された前記
製品属性パラメータの値が設定基準を満たさない場合に、前記非標準製品情報に対する前
記非製品属性パラメータの相関性が第 3 の閾値を満たすようなサブセットを決定するよう
に構成されている、システム。

【請求項 10】

請求項 9 に記載のシステムであって、
前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された前記
製品属性パラメータの少なくとも一部の値が設定基準を満たす場合に、前記少なくとも一
部の前記製品属性パラメータの値に基づいて、前記非標準製品情報が属するセットを決定
するように構成されている、システム。 40

【請求項 11】

請求項 10 に記載のシステムであって、
前記プロセッサは、更に、前記非標準製品情報を、前記非標準製品情報が属する前記セ
ットのサブセットに分配するように構成されている、システム。

【請求項 12】

請求項 11 に記載のシステムであって、 50

前記プロセッサは、更に、標準製品情報及び非標準製品情報の重複を打ち消すように構成されている、システム。

【請求項 1 3】

データ処理方法であって、

コンピュータが、複数の階層レベルを含む製品カテゴリツリーにアクセスする工程であつて、

前記複数の階層レベルのリーフカテゴリレベルは、リーフカテゴリノードを含み、

前記リーフカテゴリノードは、製品情報を含み、

前記製品情報は、複数の製品属性パラメータと、複数の非製品属性パラメータと、を含み、

前記複数の製品属性パラメータおよび前記複数の非製品属性パラメータは、標準製品情報を構成する工程と、

コンピュータが、前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択する工程と、

コンピュータが、それぞれ前記標準製品情報の少なくとも一部を含む複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割する工程と、

コンピュータが、前記複数の非製品属性パラメータの類似度を決定する工程と、

コンピュータが、前記複数のセットのうちの1つのセット内の前記標準製品情報を、各サブセット内の前記類似度が閾値を満たすように複数のサブセットに分割する工程と、

を備える方法。

【請求項 1 4】

コンピュータプログラムであって、

複数の階層レベルを含む製品カテゴリツリーにアクセスするための機能と、

前記複数の階層レベルのリーフカテゴリレベルは、リーフカテゴリノードを含み、

前記リーフカテゴリノードは、製品情報を含み、

前記製品情報は、複数の製品属性パラメータと、複数の非製品属性パラメータと、を含み、

前記複数の製品属性パラメータおよび前記複数の非製品属性パラメータは、標準製品情報を構成し、

前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択するための機能と、

それぞれ前記標準製品情報の少なくとも一部を含む複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割するための機能と、

前記複数の非製品属性パラメータの類似度を決定するための機能と、

前記複数のセットのうちの1つのセット内の前記標準製品情報を、各サブセット内の前記類似度が閾値を満たすように複数のサブセットに分割するための機能と、

をコンピュータによって実現させるコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

【関連出願の相互参照】

本出願は、あらゆる目的のために、参照によって本明細書に組み込まれる発明の名称を「AN INFORMATION PROCESSING METHOD AND SYSTEM(情報処理の方法及びシステム)」とする2010年9月21日に出願された中国特許出願第201010292384.0号に基づく優先権を主張する。

【0 0 0 2】

本開示は、コンピュータ技術の分野に関し、特に、情報処理の方法及びシステムに関する。

10

20

30

40

50

【背景技術】**【0003】**

コンピュータ技術及び通信技術の継続的発展に伴って、ますます多くのユーザが、オンラインショッピングのウェブサイト上で製品を検索、探索、及び購入するようになっている。製品の検索、探索、及び購入に先立ち、買い手ユーザは、売り手ユーザによってウェブサイト上に投稿された製品情報を閲覧することが多い。売り手ユーザは、企業体、製造業者、又は個人事業者であるかもしれない。

【0004】

売り手ユーザによってオンライン市場のウェブサイトサーバ上に掲示される製品情報は、大量である可能性がある。受信された製品情報によって表される製品を分類し、買い手ユーザを誘導して所望の情報を得られるようにするために、ウェブサイトサーバは、多階層カテゴリを使用して製品情報を分割するのが一般的である。多階層カテゴリシステムは、通常、以下で挙げられる特性を有する。

10

【0005】

第1に、多階層カテゴリシステムアーキテクチャは、比較的安定している。比較的安定したアーキテクチャを有する多階層カテゴリシステムは、売り手ユーザをシステムに慣れやすくする。このようなシステムでは、製品情報は、ウェブサイトサーバ上に掲示されるときに、多階層カテゴリシステムの要求にしたがった標準のフォーマット及びコンテンツで掲示される。このようなシステムは、また、買い手ユーザもシステムに慣れやすくし、したがって、買い手ユーザは、所望の製品情報を素早く検索する経験を蓄積することができる。

20

【0006】

第2に、多階層カテゴリシステムは、通常、ウェブサイトサーバのオペレータによって手動で操作及び維持されている。手動の操作方法によって、その分野における共通知識を規則化して標準を作ることができ、これは、多階層カテゴリシステムの使用を様々なウェブサイトに拡張するのに役立つことができる。

【0007】

上述の多階層カテゴリシステムの2つの特性を維持するためには、多階層カテゴリを使用して大量の製品情報を分割するときに、それらの製品情報を比較的粗い細分度でしか分割できないのが一般的である。なぜならば、製品情報の多様性ゆえに、製品情報が細かい細分度で分割されると、多階層カテゴリのボトム層（リーフカテゴリとも呼ばれる）は、製品情報が変わるために伴って変更されなければならず、これは、多階層カテゴリシステムの安定性にとって有害であるからであり、更に、製品情報が細かい細分度で分割されると、結果得られる多階層カテゴリアーキテクチャは、必然的に巨大になり、これは、ウェブサイトサーバの手動操作を更に難しくするからである。

30

【0008】

例えば、特定のリーフカテゴリ下の製品情報が「ドレス」であり、特定のドレスに関する情報が関心を持たれているとする。製品情報に関連付けられた製品素材がシルクからコットンに変わると、ドレスについての情報は、ドレスのリーフカテゴリ下に留まり、リーフカテゴリに対して変更を加える必要はない。しかしながら、もし、多階層カテゴリシステムの分割の細分度が更に細かく、特定のリーフカテゴリ下の製品情報が「シルクのドレス」であるならば、特定のドレスについての製品素材情報がコットンに変わると、製品情報をシルクのドレスのリーフカテゴリからコットンのドレスのリーフカテゴリに切り替える必要がある。すなわち、製品情報が変わるために伴って、リーフカテゴリが変更される。同時に、多階層カテゴリシステムは、ツリーノードアーキテクチャを使用しているのが一般的なので、サブカテゴリが追加されるたびに、多量のカテゴリが多階層カテゴリシステムに追加され、これは、多階層カテゴリシステムのアーキテクチャを非常に大きくする。

40

【0009】

多階層カテゴリシステムにおける製品情報分割の細分度は、比較的細かいので、多階層

50

カテゴリシステムの最下層に内包される製品情報の量ですら、大量である可能性がある。これらの条件下で、買い物手ユーザが多階層カテゴリシステムを使用して製品情報を検索又は照会するときは、照会時間が長びく可能性があるうえに、照会精度も低く、更に、ウェブサイトサーバが買い物手ユーザに製品情報をお勧めするときに、それをリーフカテゴリレベルでしか行えないことが多く、その結果、お勧めされる製品情報が大幅に異なり、お勧めの精度が買い物手ユーザの実際のニーズに合わなくなる。

【0010】

また、リーフカテゴリに内包される製品情報の多量さゆえに、同じリーフカテゴリに含まれる製品情報の差もまた大きく、したがって、リーフカテゴリ下で製品情報に関する作業を実現する困難度も高い。例えば、あるウェブサイトオペレータが、価格パラメータを使用して、危険な製品／偽造品についての情報を識別したいとする。通常、極めて低い価格は、偽造品である疑いが非常に強いと認識されている。オペレータは、ブランドAの製品が\$100未満であれば、その製品は偽造品の可能性が高く、種類は同じだがブランドBの製品である別の製品については、価格が\$20未満であれば、その製品は偽造品の可能性が高いというルールを設けることができる。特定の製品の価格が\$50であると、その製品が偽造品であるかどうかを価格パラメータルールを使用して直接評価することが、更に困難になる。その信頼性状態を決定するためには、製品についての更なる情報が必要になると考えられる。実際、リーフカテゴリ下の製品情報の多さは、結果として作業を極めて多量にする。

【図面の簡単な説明】

【0011】

発明の様々な実施形態が、以下の詳細な説明及び添付の図面で開示される。

【0012】

【図1】多階層カテゴリシステムの一実施形態を示す図である。

【0013】

【図2A】リーフカテゴリのなかの製品情報をサブセットに分割するためのプロセスの一実施形態を示すフローチャートである。

【図2B】リーフカテゴリのなかの製品情報をサブセットに分割するためのプロセスの一実施形態を示すフローチャートである。

【0014】

【図3】情報処理システムの一実施形態を示すブロック図である。

【0015】

【図4】標準パラメータ決定モジュールの一実施形態を示すブロック図である。

【発明を実施するための形態】

【0016】

発明は、プロセス、装置、システム、合成物、コンピュータによって読み取り可能なストレージ媒体に実装されたコンピュータプログラム製品、並びに／又は結合先のメモリに記憶された命令及び／若しくは結合先のメモリによって提供される命令を実行するように構成されたプロセッサのようなプロセッサなどの、数々の形態で実現することができる。本明細書では、これらの実現形態、又は発明がとりうるその他のあらゆる形態を、技術と称することができる。総じて、開示されるプロセスのステップの順序は、発明の範囲内で可変であることができる。別途明記されない限り、タスクを実施するように構成されるものとして説明されるプロセッサ又はメモリなどの構成要素は、所定時にタスクを実施するように一時的に構成される汎用の構成要素として、又はタスクを実施するように製造された特殊な構成要素として実装することができる。本明細書で使用される「プロセッサ」という用語は、コンピュータプログラム命令などのデータを処理するように構成された1つ以上のデバイス、回路、及び／又は処理コアを言う。

【0017】

発明の原理を例示している添付の図面とともに、以下で、発明の1つ以上の実施形態の詳細な説明が提供される。発明は、このような実施形態との関連で説明されているが、い

10

20

30

40

50

かかる実施形態にも限定されない。発明の範囲は、特許請求の範囲によってのみ限定され、発明は、数々の代替形態、変更形態、及び均等物を網羅している。以下の説明では、発明の完全な理解を与えるために、数々の具体的詳細が明記されている。これらの詳細は、例示を目的として提供されるものであり、発明は、これらの詳細の一部又は全部を伴わずとも、特許請求の範囲にしたがって実施することが可能である。明瞭さを期するために、発明に関する技術分野で知られる技術要素は、発明が必要に不明瞭にされないように、詳細な説明を省略されている。

【0018】

多階層カテゴリ構造のリーフカテゴリの動的分割が説明される。一部の実施形態では、製品カテゴリツリーのリーフカテゴリが処理され、ここでは、各リーフカテゴリについて、製品情報を代表する1つ以上の代表的な製品属性パラメータが選択される。各リーフカテゴリは、代表的な製品属性パラメータに基づいて、より細かい製品情報セットに分割される。一部の実施形態では、製品情報セットは、随意として、更にサブセットに分割される。セット及び／又はサブセットは、製品のお勧め又はユーザによる検索に使用することができる。追加の分割は、検索及び照会の時間を効果的に短縮し、検索及び照会の精度を向上させ、ウェブサイトサーバによって買い手ユーザにお勧めされる製品情報の精度を向上させる。検索、照会、及び製品お勧めの機能をサポートするための計算量もまた、軽減される。更に、この技術は、もとからのデータ構造に大幅な変更を加える必要なく標準の多階層カテゴリ構造上で動作することができる。

【0019】

図1は、多階層カテゴリシステムの一実施形態を示した図である。この例では、製品情報は、カテゴリツリーデータ構造100の形に編成されている。一部の実施形態では、カテゴリツリーコード構造100は、非一過性のストレージ媒体に記憶され、更なる処理のために、情報処理システム150によってアクセスされる。

【0020】

カテゴリツリーコード構造の上部152は、代表的な多階層カテゴリ構造である。カテゴリツリーコード構造のルートノードカテゴリ下には、N個のサブカテゴリノードを含む更に低階層のレベルがある（Nは整数である）。各サブカテゴリ下には、更なる階層レベルのノードがあってよく、その先も同様である。最下階層のレベルは、リーフカテゴリレベルとも呼ばれ、リーフカテゴリに対応するリーフノードを含む。カテゴリツリーコード構造のこの部分は、手動によるカテゴリ分けや、既存の方式に基づくカテゴリ分けなどの、既存のカテゴリ分け技術を使用して生成することができる。リーフカテゴリのなかの製品情報は、細分度が比較的粗く、多階層カテゴリ構造は、比較的安定している。

【0021】

多階層カテゴリシステム下の各リーフカテゴリは、比較的独立した安定したカテゴリであり、頻繁に変更されない。情報処理システムは、リーフカテゴリのなかの製品情報を動的に処理し、リーフカテゴリ下の1つ以上の層の形に編成される更に細かく分割された製品情報のセット及びサブセットを得る。製品情報及びセット／サブセットの分割の詳細は、後述される。一部の実施形態では、分散技術が使用され、この場合は、各リーフカテゴリは、計算ノードとして扱われ、情報処理を加速させるために、（同じコンピューティングデバイス内又は別々のデバイス内にある）プロセッサ群を使用して複数のリーフカテゴリに対して分散及び／又は並列処理作業が実施される。図1では、例えば、情報処理システムは、多階層カテゴリ構造が記憶されている環境とは別であるが、更なる処理のために、多階層カテゴリ構造にアクセスすることが可能である。

【0022】

各リーフカテゴリは、特定の製品情報に対応している。例えば、リーフカテゴリが「ドレス」であるならば、そのリーフカテゴリ下の製品情報は、ドレスに関するものでなければならない。リーフカテゴリ下の製品情報の製品属性パラメータの値が特定の設定基準を満たすかどうかに応じ、情報は、設定基準を満たす標準製品情報と、設定基準を満たさない非標準製品情報とに分割することができる。一部の実施形態では、設定基準は、製品属

10

20

30

40

50

性パラメータ値が、実際に可能な値に対応していることを要求する。例えば、設定基準は、価格帯パラメータ値が正の数値範囲であることを要求するだろう。リーフカテゴリ「ドレス」において、\$100～\$150の範囲に入る製品情報1の価格帯パラメータ値は、実際に利用可能な価格帯に対応しており、設定基準を満たしている。この価格帯は、したがって、標準製品情報として扱われる。反対に、製品情報2における価格帯パラメータ値「ABC」は、実際の価格帯に対応しておらず、これは、売り手ユーザが製品2における価格帯数値を記入したときに、エラーが生じたであろうことを意味する。したがって、製品情報2は、非標準製品情報として扱われる。

【0023】

以下の例では、リーフカテゴリ「ドレス」下に1000個の標準製品情報がある。3つの代表的な標準製品情報（標準製品情報1、標準製品情報2、及び標準製品情報3）が、表1に詳細に示されている。各標準製品情報は、4つの製品属性パラメータ、すなわちドレスの価格、ドレスの素材、ドレスのブランド、及びドレスの型番を含む。

【表1】

標準製品情報	価格	素材	ブランド	型番
標準製品情報1	価格帯1 (\$100～\$150)	素材1 (シルク)	A	1
標準製品情報2	価格値1 (\$100～\$150)	素材1 (シルク)	B	2
標準製品情報3	価格帯2 (\$50～\$100)	素材2 (コットン)	C	3

表1

【0024】

表1における4つの製品属性パラメータが、全て、リーフカテゴリ下の情報を更に細かい製品情報に分割するための分割基準として使用されたならば、分割後に得られるセットの数は、考えられる分割基準の組み合わせの数の多さゆえに、多量になると考えられる。例えば、リーフカテゴリ下の1000個の標準製品情報が、4つの異なるドレス価格帯、3種類のドレス素材、50のドレスブランド、及び5つのドレス型番に収まるならば、情報の分割は、最大で $4 \times 3 \times 50 \times 5 = 3000$ 個のセットを生み出すと考えられる。実際の処理過程では、更に多くの種類の製品属性パラメータがあるかもしれない。したがって、全ての製品属性パラメータが分割基準として利用されるならば、作業の量は、相当に大きくなる。この結果、以下で更に詳しく説明されるように、全ての製品属性パラメータのなかから一部の製品属性パラメータが、製品情報を細かく分割する目的に叶う製品属性パラメータとして選択される。

【0025】

製品情報は、製品属性パラメータと、非製品属性パラメータとを含む。本明細書で使用される製品属性パラメータとは、製品自体に固有な固定情報を言い、非限定的な例として、製品の価格、製品の素材、製品のブランド、製品の型番、及び製品の重量などが挙げられる。非製品属性パラメータは、売り手ユーザ又はウェブサイトサーバによって定められた製品に関する情報を言い、したがって、変更される可能性がある。非製品属性パラメータの例として、製品情報が属するリーフカテゴリ、投稿された製品の見出し、売り手ユーザID、製品の用途などが挙げられる。

【0026】

10

20

30

40

50

図2A～2Bは、リーフカテゴリの中の製品情報をサブセットに分割するためのプロセスの一実施形態を示したフローチャートである。プロセス100は、情報処理システム150上で実施することができる。

【0027】

ステップ101では、リーフカテゴリ下の製品情報のなかの1つ以上の製品属性パラメータが、そのリーフカテゴリ下の製品情報をサブセットに分割するために選択される。

【0028】

一部の実施形態では、製品情報を代表する1つ以上の製品属性パラメータが選択される。例えば、検索及び照会において買い手ユーザによってよく使用されるパラメータが、製品情報を代表するものと見なされる。なぜならば、選択された製品属性パラメータに基づく分割の結果として得られるセットは、ユーザによる検索及び照会の精度を向上させることができるものであるからである。更に、続くセット分割の際に可能値が多量にある結果として計算量が過剰になる問題を最小限に抑えるためには、選択された製品属性パラメータの値は、離散型である（すなわち、とりうる値が限られる）ことが望ましい。10

【0029】

これらの要求を満足させるために、一部の実施形態では、セットの分割用の製品属性パラメータを選択するための基準は、製品属性パラメータが過去の一定期間内に検索基準として使用されていること、製品属性パラメータがリーフカテゴリ下の製品情報を検索するために使用された回数が第1の閾値に達していること、及び上記製品属性パラメータが離散型であることを要求する。20

【0030】

上述の基準にしたがって、続くセット分割作業に使用するための1つ以上の製品属性パラメータが選択される。表1に示された製品情報用の製品属性パラメータの例を使用し、リーフカテゴリ下の1000個の標準製品情報が、4つのドレス価格帯、3種類のドレス素材、50のドレスブランド、及び5つのドレス型番に収まるならば、これらの値は、全て、離散型として扱うことができる。買い手ユーザが、製品情報の検索及び照会のための検索基準として価格及び素材を最も頻繁に使用し、ドレスのブランド及び型番を滅多に検索に使用せず、一定期間内に価格及び素材が検索基準として使用される回数が特定の閾値に達しているならば、価格及び素材は、このステップにおけるセット分割の目的に叶う製品属性パラメータとして決定される。30

【0031】

ステップ102では、複数の製品情報セットを得るために、選択された代表的な製品属性パラメータを使用してリーフカテゴリ下の標準製品情報が分割される。分割されたセットの中では、標準製品情報に含まれる少なくとも1つの製品属性パラメータの値が、そのセットに関連付けられた全ての製品について同一である。

【0032】

上記の例を使用し、標準製品情報を複数のセットに分割するために使用される製品属性パラメータが、4つの製品価格帯及び3種類の製品素材であるならば、分割後、以下の $4 \times 3 = 12$ 個の組み合わせの製品データセットが得られる。

【0033】

- セット1：価格帯1 + 素材1
- セット2：価格帯1 + 素材2
- セット3：価格帯1 + 素材3
- セット4：価格帯2 + 素材1
- セット5：価格帯2 + 素材2
- セット6：価格帯2 + 素材3
- セット7：価格帯3 + 素材1
- セット8：価格帯3 + 素材2
- セット9：価格帯3 + 素材3
- セット10：価格帯4 + 素材1

1020304050

セット11：価格帯4+素材2

セット12：価格帯4+素材3

【0034】

同じ分割セットに属するどの2つの標準製品情報に含まれる価格帯及び素材も、同一である。例として、表1に示された標準製品情報用の製品属性パラメータを使用すると、標準製品情報1及び標準製品情報2は、セット1に分配されるべきであり、標準製品情報3は、セット5に分配されるべきである。

【0035】

ステップ101において、製品情報の細かい分割に使用される製品属性パラメータの選択の際は、買い物手ユーザによる検索及び照会の習慣が十分に考慮され、製品属性パラメータ値が離散型である一方で、選択されなかった製品属性パラメータは放置されるので、このステップによる分割結果は、買い物手ユーザの利用習慣を正確に反映しているのみならず、分割の際の作業の量も軽減する。

10

【0036】

上記のステップ101及び102の使用を通じて、多階層カテゴリシステム下のリーフカテゴリのなかの製品情報は、更に細かい分割製品情報セットを得るために自動的に分割され、したがって、買い物手ユーザが検索又は照会を行うときは、更に細かい製品セットが検索される。これは、待ち時間を効果的に短縮し、全体的な精度及びウェブサイトサーバによって買い物手ユーザにお勧めされる製品情報の精度を向上させ、製品情報に対して実施される様々な作業の実行可能性の向上を助ける。例えば、ドレスのリーフカテゴリが12個のセットに分割された後は、買い物手ユーザの検索基準に基づく検索を、対応するセットの中で実施することができる。各セットにおける製品情報の量は、リーフカテゴリ下の製品情報の量よりも大幅に少ないので、検索の際の買い物手ユーザの待ち時間が大幅に短縮され、検索及び照会の精度が向上される。それと同時に、買い物手ユーザに製品情報がお勧めされるときに、特定の製品属性（例えば指定の価格レベル及び特定の種類の素材）に基づいて買い物手ユーザへの製品情報のお勧めがなされることによって、勧められる製品情報が買い物手ユーザの実際のニーズに近づくことが可能になる。また、引き続き、安全でない製品についての情報をとるために価格パラメータを使用する例において、上述の方法を使用して製品情報が細かく分割された後は、探されている価格に基づいて、対応するセットを素早く見つけることができ、次いで、見つけられたセットの中で、偽造品の存在をチェックするための照会を実施することができ、そうして、作業に関連した計算の量を効率的に軽減し、実行の有効性を高めることができる。

20

【0037】

セット分割のために、製品情報を最も良く反映する離散型の製品属性パラメータ値を使用することによって、分割後に得られたセットに基づくユーザによる検索及び照会の精度を向上させることができる。そして、標準製品情報に含まれる製品属性パラメータの可能値が過剰にある結果としてセット分割の際の作業の量が多量になる問題も、最小限に抑えることができる。

30

【0038】

一部の実施形態では、随意としてステップ103において、各分割セットについて、セット内の全ての標準製品情報に含まれる1つ以上の非製品属性パラメータにしたがってセット内の標準製品情報が更に複数のサブセットに分割される。

40

【0039】

一部の実施形態では、セット及び／又はサブセットに含まれる処理済みの製品情報は、買い物手ユーザに製品のお勧めをするために及び／又は買い物手ユーザによる製品検索を助けるために使用される。例えば、セット及び／又はサブセットのなかの製品情報は、買い物手ユーザの要求（買い物手ユーザによって入力されたキーワード又は観測された買い物手ユーザの購入習慣）に基づいて買い物手ユーザに送信することができる。別の例としては、検索エンジンがセット及び／又はサブセットのなかでキーワード検索を実施し、その結果を買い物手ユーザに送ることが可能である。

50

【0040】

引き続き、ドレスに関する上記の例について、標準製品情報1（例えば価格帯\$100-\$150、シルク素材、ブランドA、型番1）及び標準製品情報2（例えば価格帯\$100-\$150、シルク素材、ブランドB、型番2）が、セット1に分配され、製品情報の非製品属性パラメータが、売り手ユーザによってウェブサイトサーバ上に投稿された製品ヘッダであると想定する。例えば、標準製品情報1内に投稿された製品ヘッダのキーワードは、「ドレス」及び「アニマル柄」であり、標準製品情報2内に投稿された製品ヘッダのキーワードは、「ドレス」及び「ストライプ」である。

【0041】

一部の実施形態では、次に、同じセット内の標準製品情報の非製品属性の類似度が決定される。類似度の計算には、様々な標準的技術を使用することができる。例えば、一部の実施形態では、セット内の非製品属性の類似度を計算するために、K-medoidsアルゴリズムが使用される。

10

【0042】

最後に、1つのサブセット内の、ペアをなす製品情報の非製品属性パラメータの類似度が第2の閾値を満たすように、同じセット内の全ての標準製品情報が、更に複数のサブセットに分割される。一部の実施形態では、類似度を決定するために、K-medoidsなどの技術が使用される。類似度が第2の閾値を満たす製品情報のみが同じサブセットに分配されるので、異なるサブセット間の非製品属性パラメータの類似度は、比較的低いと見なされる。

20

【0043】

本実施形態は、また、製品情報のその他の非製品属性パラメータを使用したサブセット分割に限定されず、製品の用途に基づく分割などがある。

【0044】

標準製品情報の細かい分割は、上記のステップ101～103の提示内容の使用によって達成される。更に、非標準製品情報を更に細かいセットに分割することも可能である。

【0045】

ステップ104では、リーフカテゴリ下の非標準製品情報が決定される。

【0046】

非標準製品情報の製品属性パラメータの一部又は全部の値は、設定基準を満たさないが、もし、これらの値が、情報を掲示するときに売り手ユーザ側で生じた完了エラー又はその他の人為的エラーの結果であるならば、非標準製品情報は、依然として製品に関する関連情報を表すことができ、検索、照会、ユーザへのお勧め、又はその他の作業での使用にとっても大きな意味を持つ。

30

【0047】

ステップ105では、非標準製品情報に含まれる、セット分割に使用された製品属性パラメータの全ての値が特定の設定基準（例えば値のフォーマットや値の範囲など）を満たさないか否かについて、決定が下される。満たさないならば、制御は、ステップ106に引き渡され、満たすならば、制御は、ステップ108に引き渡される。

【0048】

40

非標準製品情報に含まれる、セット分割に使用された製品属性パラメータの全ての値が設定基準を満たさないかどうかを決定するために、一部の実施形態では、セット分割における使用のためにステップ101において選択された製品属性パラメータが設定基準を満たすパラメータであるか否かが決定される。例えば、ステップ101において、セット分割における使用のために選択された製品属性パラメータが製品価格及び製品素材であるならば、ステップ105では、非標準製品情報に含まれる2つの製品属性パラメータ、すなわち製品価格及び製品素材がともに設定基準を満たさないかどうかが決定される。

【0049】

ステップ101における製品情報の全ての製品属性パラメータがセット分割に使用されたならば、設定基準を満たさない値の製品属性パラメータを有する製品情報がある限り、

50

そのような情報は、非標準製品情報であると見なされる。しかしながら、ステップ101において、製品属性パラメータの一部のみがセット分割における使用のために選択されたならば、設定基準を満たさないその一部の製品属性パラメータの値のみが非標準製品情報であると見なされる。

【0050】

具体的には、設定基準を満たさない値の製品属性パラメータを含む特定の非標準製品情報があるならば、その非標準製品情報は、セット分割には使用されず、標準製品情報の特殊なケースとして扱われ、ステップ102において分割を経ることができる。あるいは、非標準製品情報は、依然として非標準製品情報として扱われ、後続ステップにおいて分割を経ることもできる。例えば、セット分割に使用された製品属性パラメータが製品価格及び製品素材であり、しかしながら、製品型番の製品属性パラメータが設定基準を満たさないならば、製品型番は、標準製品情報の特殊なケースとして扱われ、依然としてステップ102にしたがって分割されるか、或いは非標準製品情報として扱われ、後述される後続ステップにおいて分割されるかのいずれかであることができる。10

【0051】

ステップ106では、非標準製品情報に含まれる少なくとも1つの非製品属性パラメータが決定され、非標準製品情報に対する非製品属性パラメータの相関性が第3の閾値を満たすようなサブセットが存在するかどうかが決定される。存在するならば、非標準製品情報は、第3の閾値を満たす相関性が最も高いサブセットに分配され(ステップ113)、制御は、ステップ111に引き渡され、もし存在しないならば、プロセスは、ステップ107に続く。20

【0052】

非標準製品情報の製品属性パラメータは、セット分割に使用することができないので、ステップ106では、非標準製品情報が属するサブセットを決定するために、非標準製品情報の非製品属性パラメータが使用される。

【0053】

引き続き、製品価格及び製品素材をセット分割のための製品属性パラメータとして、投稿された製品ヘッダをサブセット分割に使用される非製品属性パラメータとして使用する例について、ステップ106において、特定の非標準製品情報の製品価格及び製品素材の値がともに設定基準を満たさないならば、その非標準製品情報の投稿製品ヘッダ(非製品属性パラメータ)が決定され、K-medoidsアルゴリズムなどのクラスタリングアルゴリズムを使用して、各サブセットに含まれる全ての標準製品情報の投稿製品ヘッダに対するその投稿製品ヘッダの類似度が計算される。作業の量を軽減するためには、各サブセットのなかの一標準製品情報の投稿製品ヘッダに対して同様の計算を実施して、第3の閾値を満たす標準製品情報が存在するサブセットを特定することができ、非標準製品情報は、類似度が最も高い標準製品情報が存在するサブセットに分配することができる。30

【0054】

ステップ107では、特別に定められたセットに非標準製品情報が分配され、プロセスは、ステップ111に進む。ステップ107では、一部の非標準情報のセット分割に使用されたが設定基準を満たさない製品属性パラメータの値の存在、及びどのセットの中のサブセットとも相關しない非製品属性パラメータの存在ゆえに、このタイプの非標準製品情報のために、特別に定められた別のセットが設定される。この特別に定められたセット及びステップ102において分割されたセットは、クラス間の非類似性の原則を満足している。

【0055】

ステップ108では、非標準製品情報に含まれる製品属性パラメータの中で、設定基準を満たす値を持つ製品属性パラメータが決定される。この時点では、非標準製品情報の、セット分割に使用された製品属性パラメータの一部の値が設定基準を満たしているので、非標準製品情報の分割には、設定基準を満たす値を有する上記一部の製品属性パラメータを使用することができる。4050

【0056】

例えば、セット分割における使用のためにステップ101において決定された製品属性パラメータが製品価格及び製品素材であり、或る非標準製品情報の製品価格の値が設定基準を満たさず、しかしながら、製品素材の値は設定基準を満たすならば、非標準製品情報の分割には、製品属性パラメータ「製品素材」を使用することができる。

【0057】

ステップ109では、ステップ108において決定された製品属性パラメータに基づいて、非標準製品情報が属するセットが決定される。この時点では、決定された製品属性パラメータの値と、各セットにおける標準製品情報に含まれる製品属性パラメータの値とに基づいて、非標準製品情報の製品属性パラメータの値と同じ値を持つセットが特定される

10

。

【0058】

決定されるセットの数は、2つ以上であってよく、このような状況下では、決定された全てのセットに対して後続ステップを実行することができる。例えば、ステップ101に関連して例示された例では、12個のセットが生成された。もし、このステップにおける製品属性パラメータが、製品素材値1を持つ製品素材であると想定されるならば、決定された非標準製品情報が属するセットは、セット1、セット4、セット7、又はセット10であることが可能である。

【0059】

ステップ110では、非標準製品情報の非製品属性パラメータに基づいて、非標準製品情報が、ステップ109において事前に決定されたセットのサブセットに分配される。

20

【0060】

ステップ109の後は、非標準製品情報が属するセットが決定されが、どのセットのどのサブセットに非標準製品情報が属するかは決定されていない。したがって、ステップ109は、更に、非製品属性パラメータを利用し、類似度が最も高い非製品属性パラメータを有するサブセットに非標準製品情報を分配することによって、非標準製品情報の更なる分割を実施する。

【0061】

ステップ106において、非標準製品情報が分配されたときは、その分配先のサブセットに対する類似度が、第3の閾値を満たす必要があり、そして、類似度が最も高いサブセットが選択された。これに対し、ステップ109では、ステップ109において確認されたセットの中から、類似度が最も高いサブセットを選択するだけでよい。なぜならば、ステップ105では、非標準製品情報のなかの、セット分割に使用された製品属性パラメータのいずれも、設定基準を満たさなかったからである。したがって、最小閾値（すなわち第3の閾値）が設定されず、単に、非標準製品情報の非製品属性パラメータに対する相関性が最も高いサブセットが選択されたならば、どのサブセットに対しても、全ての非標準製品情報の相関性が非常に低くなる可能性がある。しかしながら、そのサブセットは、依然として非標準製品情報が属するサブセットとして選択され、その結果、非標準製品情報と、同じサブセットのなかの他の標準製品情報との間の相関性は、低くなると考えられる。しかしながら、ステップ109において決定されたセットは、設定基準を満たす値を有する1つの非標準製品情報製品属性パラメータを使用して決定されたので、プロセスのステップ110では、ステップ110において選択に使用されたセットと、非標準製品情報との間に幾らかの相関性があり、したがって、そこから選択されたサブセットと、非標準製品情報との間の相関性もまた、比較的高くなる。

30

【0062】

売り手ユーザによってウェブサイトサーバ上に投稿される製品情報は、重複する可能性がある。重複する製品情報の処理に必要とされるリソースを減らすために、ステップ111は、標準製品情報及び非標準製品情報に対して重複打ち消し操作を実施する。具体的には、ステップ111において、重複する標準製品情報が決定され、そのような標準製品情報の1つが維持され、残りの同一の標準製品情報は排除される。更に、重複する非標準製

40

50

品情報が決定され、そのような非標準製品情報の 1 つが維持され、残りの同一の非標準製品情報は排除される。

【 0 0 6 3 】

2 つの標準製品情報の製品属性パラメータ及び非製品属性パラメータが同一であるならば、これら 2 つの標準製品情報は、重複する標準製品情報として定義される。

【 0 0 6 4 】

2 つの非標準製品情報の製品属性パラメータ及び非製品属性パラメータが同一であるならば、これら 2 つの非標準製品情報は、重複する非標準製品情報として定義される。

【 0 0 6 5 】

ステップ 111 の重複打ち消し操作は、ステップ 101 の前に実行することができる、
又はステップ 101 ~ 110 の間の任意で実行することができる、又はステップ 110 の
後に実行することができる。
10

【 0 0 6 6 】

ステップ 112 において、各サブセットは、名前すなわちラベルを割り当てられる。ステップ 112 は、ステップ 110 の後、ステップ 111 の前に実行することもできる。ステップ 110 を経て実行された時点で、製品情報は、実際は、既に、標準製品ユニット (S P U) を得るために自動的に分割されており、これらのユニットは、セットレベル又はサブセットレベルにある。S P U の再利用を促すために、プロセスは、各 S P U に名前又はラベルを割り当て、これは、各 S P U によって表されるコンテンツも表している。

【 0 0 6 7 】

具体的には、一部の実施形態では、先ず、各サブセットの標準製品情報及び非標準製品情報の製品属性パラメータ及び非製品属性パラメータが決定される。次いで、上記の製品属性パラメータ及び非製品属性パラメータに含まれる少なくとも 1 つの高頻度ワードが決定される。最後に、上記の高頻度ワードが、上記サブセットの名前として割り当てられる。
20

【 0 0 6 8 】

続く買い手ユーザによる検索及び照会の最中、並びに買い手ユーザに製品情報をお勧めする際に、サブセットの名前に関わっている高頻度ワードは、検索又はお勧めを実施するためのキーワードの役割を果たすことができる。

【 0 0 6 9 】

図 3 は、情報処理システムの一実施形態を示したブロック図である。この例では、システム 300 は、例えば、パソコン、サーバコンピュータ、手持ち式のすなわち携帯型のデバイス、フラットパネルデバイス、マルチプロセッサシステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラム可能な家庭用電子機器、ネットワーク P C、ミニコンピュータ、大型コンピュータ、特殊用途向けデバイス、上記のシステム若しくはデバイスの任意を含む分散コンピューティング環境、又は 1 つ以上のプロセッサと該プロセッサにつながれて該プロセッサに命令を提供するように構成されたメモリとを含むその他のハードウェア / ソフトウェア / ファームウェアの組み合わせなどの、1 つ以上のコンピューティングデバイスを使用して実装することができる。
30

【 0 0 7 0 】

システム 300 は、標準パラメータ決定モジュール 11 と、第 1 の分割モジュール 12 とを含み、標準パラメータ決定モジュール 11 は、リーフカテゴリ下の全ての標準製品情報に含まれる少なくとも 1 つの製品属性パラメータを決定するために使用され、第 1 の分割モジュール 12 は、そうして決定された少なくとも 1 つの製品属性を使用して、標準製品情報を分割し、上述のリーフカテゴリ下に複数のセットを得るために使用され、ここでは、同じセットに分配された標準製品情報に含まれる上述の少なくとも 1 つの製品属性パラメータの値は、同一である。
40

【 0 0 7 1 】

図 4 は、標準パラメータ決定モジュールの一実施形態を示したブロック図である。この例では、モジュールは、属性パラメータ決定サブモジュール 21 と、頻度決定サブモジュ
50

ール 2 2 と、選択サブモジュール 2 3 とを含み、属性パラメータ決定サブモジュール 2 1 は、リーフカテゴリ下の全ての標準製品情報に含まれる全ての製品属性パラメータを決定するために使用され、頻度決定サブモジュール 2 2 は、設定期間内に、上述のリーフカテゴリ下で各製品属性パラメータを検索基準として使用した製品検索が実施される回数を決定し、選択サブモジュール 2 3 は、全ての上述の製品属性パラメータのなかから少なくとも 1 つの製品属性パラメータを選択するために使用され、ここでは、選択されたどの製品属性パラメータが製品情報検索のための検索基準として使用された回数も、第 1 の閾値を満たしており、上記製品属性パラメータの値は、離散型である。

【 0 0 7 2 】

図 3 に戻り、システム 3 0 0 は、更に、非標準パラメータ決定モジュール 1 3 と、類似性決定モジュール 1 4 と、第 2 の分割モジュール 1 5 とを含み、非標準パラメータ決定モジュール 1 3 は、各分割セットに対し、それらのセットのなかの全ての標準製品情報に含まれる少なくとも 1 つの非製品属性パラメータを決定するために使用され、類似性決定モジュール 1 4 は、同じセット内の全ての標準製品情報の少なくとも 1 つの非製品属性パラメータの類似度を決定するために使用され、第 2 の分割モジュール 1 5 は、類似度にしたがって、同じセット内の全ての標準製品情報を更に分割し、上記セットの複数のサブセットを得るために使用され、ここでは、同じサブセット内のどの 2 つの標準製品情報間の上述の少なくとも 1 つの非製品属性パラメータの類似度も、第 2 の閾値を満たしている。

【 0 0 7 3 】

上述の機器は、更に、判定モジュール 1 6 を含み、該モジュールは、非標準製品情報に含まれる、セット分割に使用された全ての製品属性パラメータの値が設定基準を満たさないかどうかを判定するために使用され、満たさないならば、非標準パラメータ決定モジュールがトリガされ、満たすならば、標準パラメータ決定モジュールがトリガされ、上述の非標準製品情報のなかの、セットの分割に使用された少なくとも 1 つの製品属性パラメータの値は、設定基準を満たす。

【 0 0 7 4 】

判定モジュール 1 6 によるトリガに基づく非標準パラメータ決定モジュール 1 3 及び標準パラメータ決定モジュール 1 1 のそれぞれの作業プロセスは、以下に挙げられるとおりである。

【 0 0 7 5 】

非標準製品情報に含まれる、セット分割に使用された全ての製品属性パラメータの値が設定基準を満たさないとき：

【 0 0 7 6 】

上述の非標準パラメータ決定モジュール 1 3 は、更に、上記非標準製品情報に含まれる少なくとも 1 つの非製品属性パラメータを決定するために使用され、上述の類似性決定モジュール 1 4 は、非標準製品情報の非製品属性パラメータと、各サブセットに含まれる標準製品情報の非製品属性パラメータとの間の類似度を決定するために使用され、上述の第 2 の分割モジュール 1 5 は、更に、非製品属性パラメータに基づいて、上記非標準製品情報をサブセットに分配するために使用され、ここでは、上記非標準製品情報の非製品属性パラメータの、その分配先のサブセット内の標準製品情報に含まれる非製品属性パラメータに対する類似度は、第 3 の閾値を満たしている。

【 0 0 7 7 】

具体的には、非標準製品情報のなかの、セット分割に使用された全ての製品属性パラメータの値が設定基準を満たさず、上記非標準製品情報の非製品属性パラメータの、どのサブセットの標準製品情報に含まれる非製品属性パラメータに対する類似度も第 3 の閾値未満であるときは、上述の第 1 の分割モジュール 1 2 は、更に、非標準製品情報のなかの、セット分割に使用された全ての製品属性パラメータの値が設定基準を満たさず、上記非標準製品情報の非製品属性パラメータの、どのサブセットの標準製品情報に含まれる非製品属性パラメータに対する類似度も第 3 の閾値未満であるときの、特別に定められたセットに、上記非標準製品情報を分配するために使用される。

10

20

30

40

50

【 0 0 7 8 】

非標準製品情報に含まれる、セット分割に使用された製品属性パラメータの値の全ては、設定基準を満たさないとき：

【 0 0 7 9 】

上述の標準パラメータ決定モジュール 1 1 は、更に、非標準製品情報のなかの、セット分割に使用された製品属性パラメータのうちの、どの製品属性パラメータが設定基準を満たす値を有するのかを決定するために使用され、上述の第 1 の分割モジュール 1 2 は、更に、設定基準を満たす値を有する上述の製品属性パラメータを内包するセットを決定するために使用され、上述の第 2 の分割モジュール 1 5 は、更に、非標準製品情報の非製品属性パラメータに基づいて、上記非標準製品情報を、第 1 の分割モジュール 1 2 によって決定されたセット内のサブセットに分配するために使用され、ここでは、上記非標準製品情報の非製品属性パラメータの、その分配先のサブセット内の標準製品情報の非製品属性パラメータに対する類似度は、最も高くなっている。

10

【 0 0 8 0 】

上述の機器は、更に、重複打ち消しモジュール 1 7 を含み、該モジュールは、重複する標準製品情報を決定し、そのような標準製品情報の 1 つを維持し、残りのそのような標準製品情報を排除するために、及び重複する非標準製品情報を決定し、そのような非標準製品情報の 1 つを維持し、残りのそのような非標準製品情報を排除するために使用される。上述の重複する標準製品情報の間では、全ての製品属性パラメータ及び非製品属性パラメータが同一であり、上述の重複する非標準製品情報の間では、全ての製品属性パラメータ及び非製品属性パラメータが同一である。

20

【 0 0 8 1 】

上述の機器は、更に、命名モジュール 1 8 を含み、該モジュールは、各サブセットの標準製品情報及び非標準製品情報に含まれる製品属性パラメータ及び非製品属性パラメータを決定するために使用される。上記の製品属性パラメータ及び非製品属性パラメータに含まれる少なくとも 1 つの高頻度ワードが決定され、該少なくとも 1 つの高頻度ワードは、上記サブセットを命名するために使用される。

【 0 0 8 2 】

本開示の実施形態で提供されている方法及び機器の使用は、買い手ユーザが細かい製品情報セットに基づいて製品情報の検索又は照会を行うときに、検索及び照会の回数が効果的に減らされ、検索及び照会の精度が向上され、ウェブサイトサーバによって買い手ユーザにお勧めされる製品情報の精度が向上されるように、製品属性を基準として使用してリーフカテゴリベースで製品情報を細かいセットに分割することを可能にし、更に、非製品属性を基準として使用して更に細かい細分度のサブセットに分割することを可能にする。また、細かい製品情報セットは、製品情報に対して実施される作業の実行可能性を向上させ、製品情報に対して作業が実施されるときの計算の量を軽減することができる。更に、多階層カテゴリシステムの実質的内容は、セット分割及びサブセット分割の過程で変更されることはなく、既存の多階層カテゴリシステムの特徴に適合したままであり、買い手ユーザの検索習慣は、十分に考慮され、セット分割は、買い手ユーザによる検索でよく使用される製品属性パラメータを用いる。また、非標準製品情報に含まれる製品属性パラメータの値が設定基準を満たさないような様々な状況に対応する分割手段が提供され、したがって、可能な限り、全ての非標準製品情報が、相関性が最も高いサブセットに分配される。最後に、本開示の提案は、製品情報の重複打ち消し及び分割サブセットの命名も実施し、重複する製品情報の処理に用いられるリソースを効果的に軽減するとともに、管理者がサブセットの名前に基づいてサブセットを管理すること及びサブセットの名前をキーワードとして使用してサブセット内の製品情報を照会することを容易にする。

30

【 0 0 8 3 】

上述されたモジュールは、1 つ以上の汎用プロセッサ上で実行されるソフトウェアコンポーネントとして、プログラマブルロジックデバイス及び / 若しくは特定の機能を実施するように設計された特殊用途向け集積回路などのハードウェアとして、又はそれらの組み

40

50

合わせとして実装することができる。一部の実施形態では、モジュールは、本発明の実施形態で説明される方法を（パソコン、サーバ、ネットワーク機器などの）コンピュータデバイスに実行させるための命令を含み尚且つ（光ディスク、フラッシュストレージデバイス、モバイルハードディスクなどの）不揮発性のストレージ媒体に記憶させることができるソフトウェア製品の形で具現化することができる。モジュールは、1つのデバイスに実装する又は複数のデバイスに分散させることができる。モジュールの機能は、互いに合体させる又は複数のサブモジュールに更に分けることができる。

【 0 0 8 4 】

当業者ならば、本開示の実施形態が、方法、システム、又はコンピュータソフトウェア製品として提供できることがわかる。したがって、本開示は、完全にハードウェアで構成される実施形態、完全にソフトウェアで構成される実施形態、又はハードウェアとソフトウェアとを組み合わせた実施形態の形態をとることができる。更に、本開示は、コンピュータプログラムコードを含む1つ以上のコンピュータによって操作可能なストレージ媒体（磁気ディスクストレージデバイス、CD-ROM、及び光ストレージデバイスを非限定例として含む）に実装されたコンピュータプログラムの形態をとることができる。10

【 0 0 8 5 】

本開示は、方法、機器（システム）、及びコンピュータプログラム製品に基づくフローチャート及び／又はブロック図を参照にして説明されている。フローチャート及び／又はブロック図における各プロセス及び／又は各ブロック、並びにフローチャート及び／又はブロック図におけるプロセス及び／又はブロックの組み合わせは、それぞれ、コンピュータプログラムコマンドを通じて達成できることがわかる。これらのコンピュータプログラムコマンドは、汎用コンピュータ、特殊用途コンピュータ、埋め込みプロセッサ、又はその他のプログラマブルデータ機器のプロセッサに提供されて、マシンを作動させることができ、その結果、コンピュータ、又はその他のプログラマブルデータ機器のプロセッサを通じて実行されるコマンドは、フローチャートの1つ以上のプロセス及び／又はブロック図の1つ以上のブロックによって指定された機能を実現するために使用されるデバイスを作動させる。20

【 0 0 8 6 】

一部の実施形態では、これらのコンピュータプログラムコマンドは、また、コンピュータ又はその他のプログラマブルデータ機器を誘導することができる特殊作業用コンピュータによって読み取り可能なストレージデバイスに記憶させることもでき、その結果、これらのコンピュータ読み取り可能ストレージデバイスに記憶されているコマンドは、コマンドデバイスを含む製品をもたらす。これらのコマンドデバイスは、フローチャートの1つ以上のプロセス及び／又はブロック図の1つ以上のブロックに指定される機能を実現する。30

【 0 0 8 7 】

これらのコンピュータプログラムコマンドは、コンピュータ又はその他のプログラマブルデータ機器に取り込むこともでき、その結果、コンピュータ又はその他のプログラマブル機器において一連の作業ステップが実行されて、コンピュータ処理が引き起こされる。このようにして、コンピュータ又はその他のプログラマブル機器において実行されるコマンドは、フローチャートの1つ以上のプロセス及び／又はブロック図の1つ以上のブロックで指定される機能を実現するために使用されるステップを提供する。40

【 0 0 8 8 】

本出願の好ましい実施形態が既に説明されてきたが、当業者ならば、ひとたび基本的な発明的概念を捉えたら、これらの実施形態に対してその他の変更又は修正を加えることができる。したがって、添付の特許請求の範囲は、好ましい実施形態のみならず、あらゆる変更及び修正も本出願の範囲内に入るものとして含むとして解釈される。

【 0 0 8 9 】

当業者ならば、本発明の趣旨及び範囲から逸脱することなく本出願を変更及び多様化することができる。したがって、もし、本出願のこれらの変更及びバリエーションが、特50

許請求の範囲及びその等価技術の範囲内であるならば、本出願は、これらの変更及びヴァリエーションも網羅することを意図される。

【 0 0 9 0 】

以上の実施形態は、理解を明瞭にする目的で幾らか詳細に説明されてきたが、発明は、提供された詳細に限定されない。発明を実現するには、多くの代替的手法がある。開示された実施形態は、例示的なものであり、限定を目的としたものではない。

適用例 1：データ処理システムであって、1つ以上のプロセッサであって、複数の階層レベルを含む製品カテゴリツリーにアクセスするように構成され、前記複数の階層レベルのリーフカテゴリツリーは、リーフカテゴリノードを含み、前記リーフカテゴリノードは、製品情報を含み、前記製品情報は、複数の製品属性パラメータを含み、前記複数の製品属性パラメータは、標準製品情報を構成し、前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択するように構成され、複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割するように構成されている、1つ以上のプロセッサと、前記複数のセットはそれぞれ前記標準製品情報の少なくとも一部を含むように構成されており、前記1つ以上のプロセッサに接続され、前記1つ以上のプロセッサに命令を提供するように構成されている1つ以上のメモリと、を備える、システム。

適用例 2：適用例 1に記載のシステムであって、前記代表的な製品属性パラメータは、製品を検索するためにユーザによって以前に使用されている、システム。

適用例 3：適用例 1に記載のシステムであって、前記複数のセットのうちの1つのセットは、前記セットに関連付けられている全ての製品において同一の値の製品属性パラメータを含む、システム。

適用例 4：適用例 1に記載のシステムであって、前記プロセッサは、更に、製品のお勧めをするために前記複数のセットを使用するように構成されている、システム。

適用例 5：適用例 1に記載のシステムであって、前記プロセッサは、更に、検索結果を提供するために前記複数のセットを検索するように構成されている、システム。

適用例 6：適用例 1に記載のシステムであって、前記標準製品情報は、複数の非製品属性パラメータを含む、システム。

適用例 7：適用例 6に記載のシステムであって、前記プロセッサは、更に、非製品属性パラメータの類似度を決定し、前記複数のセットのうちの1つのセット内の前記標準製品情報を、各サブセット内の前記類似度が閾値を満たすように複数のサブセットに分割するように構成されている、システム。

適用例 8：適用例 7に記載のシステムであって、前記プロセッサは、更に、製品のお勧めをするために前記複数のサブセットを使用するように構成されている、システム。

適用例 9：適用例 1に記載のシステムであって、前記複数の製品属性パラメータは、更に、非標準製品情報を構成する、システム。

適用例 10：適用例 7に記載のシステムであって、前記複数の製品属性パラメータは、更に、非標準製品情報を構成し、前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された製品属性パラメータの値が設定基準を満たすかどうかを決定するように構成されている、システム。

適用例 11：適用例 10に記載のシステムであって、前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された前記製品属性パラメータの値が設定基準を満たさない場合に、前記非標準製品情報に対する前記非製品属性パラメータの相関性が第3の閾値を満たすようなサブセットを決定するように構成されている、システム。

適用例 12：適用例 11に記載のシステムであって、前記プロセッサは、更に、前記非標準製品情報に対応し、セット分割に使用された前記製品属性パラメータの少なくとも一部の値が設定基準を満たす場合に、前記少なくとも一部の前記製品属性パラメータの値に基づいて、前記非標準製品情報が属するセットを決定するように構成されている、システム。

10

20

30

40

50

適用例 1 3 : 適用例 1 2 に記載のシステムであって、前記プロセッサは、更に、前記非標準製品情報を、前記非標準製品情報が属する前記セットのサブセットに分配するよう構成されている、システム。

適用例 1 4 : 適用例 1 3 に記載のシステムであって、前記プロセッサは、更に、標準製品情報及び非標準製品情報の重複を打ち消すように構成されている、システム。

適用例 1 5 : データ処理方法であって、複数の階層レベルを含む製品カテゴリツリーにアクセスし、前記複数の階層レベルのリーフカテゴリレベルは、リーフカテゴリノードを含み、前記リーフカテゴリノードは、製品情報を含み、前記製品情報は、複数の製品属性パラメータを含み、前記複数の製品属性パラメータは、標準製品情報を構成し、前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択し、複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割し、前記複数のセットは、それぞれ前記標準製品情報の少なくとも一部を含むこと、を備える方法。

10

適用例 1 6 : 有形のコンピュータによって読み取り可能なストレージ媒体に実装されているコンピュータプログラム製品であって、複数の階層レベルを含む製品カテゴリツリーにアクセスするためのコンピュータ命令と、前記複数の階層レベルのリーフカテゴリレベルは、リーフカテゴリノードを含み、前記リーフカテゴリノードは、製品情報を含み、前記製品情報は、複数の製品属性パラメータを含み、前記複数の製品属性パラメータは、標準製品情報を構成し、前記リーフカテゴリノードに対応する前記複数の製品属性パラメータの中から、前記製品情報を代表する代表的な製品属性パラメータを選択するためのコンピュータ命令と、複数のセットを得るために、前記代表的な製品属性パラメータを使用して前記リーフカテゴリノードの標準製品情報を分割するためのコンピュータ命令と、前記複数のセットは、それぞれ前記標準製品情報の少なくとも一部を含むこと、を備えるコンピュータプログラム製品。

20

【図1】

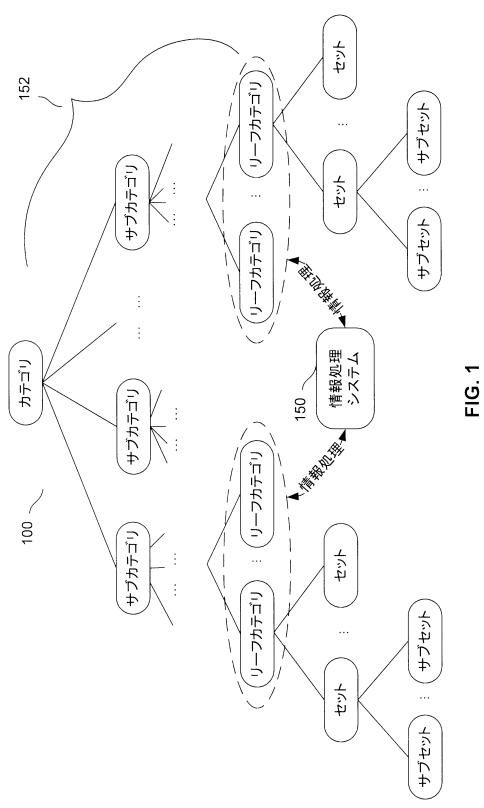


FIG. 1

【図2A】

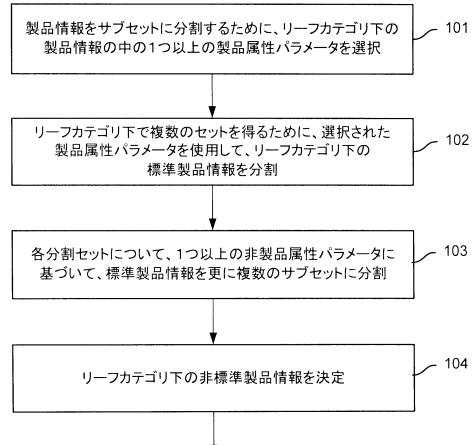


FIG. 2A

【図2B】

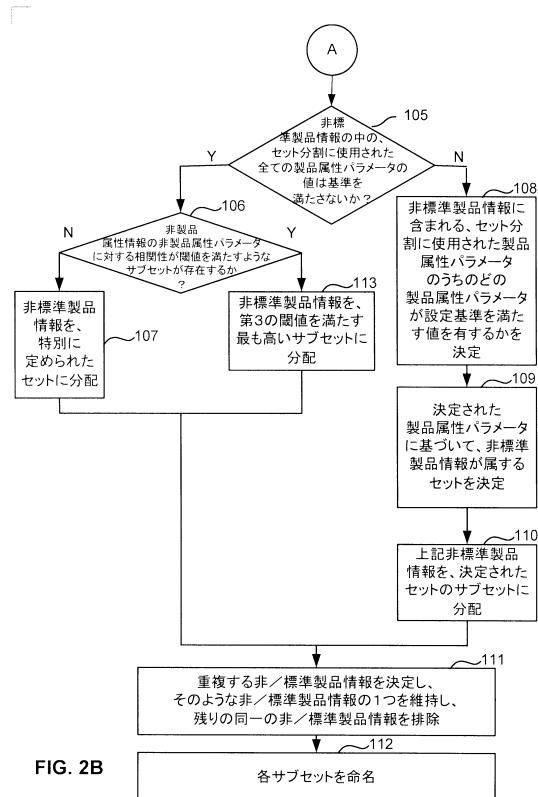


FIG. 2B

【図3】

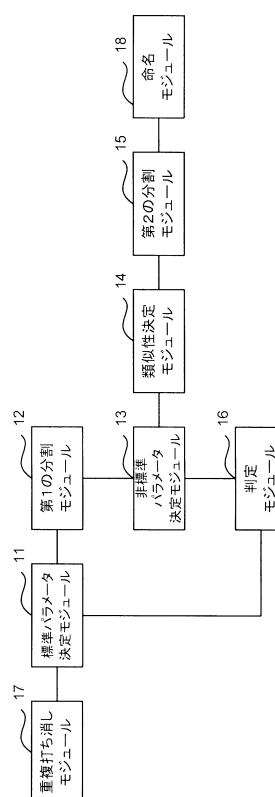


FIG. 3

【図4】

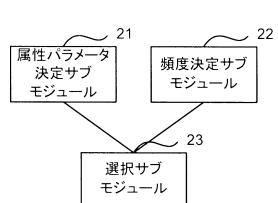


FIG. 4

フロントページの続き

(72)発明者 グウ・ハイジエ

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

(72)発明者 スウ・ニーンジュン

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

(72)発明者 ダイ・オイフォン

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

(72)発明者 マー・ハイピーン

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

(72)発明者 ジャーン・ジンイン

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

(72)発明者 チエン・エンホーン

中華人民共和国 ハンチョウ , ワーナー・ロード , ザ・ウエスト・レイク・インターナショナル・
プラザ・オブ・エス アンド ティー , ビルディング エー , 10階 , ナンバー391 , アリババ
・グループ・リーガル・デパートメント内

審査官 野崎 大進

(56)参考文献 特開2000-163494(JP,A)

特開2010-176360(JP,A)

特開2009-129087(JP,A)

特開2007-115222(JP,A)

特開2004-295279(JP,A)

濫谷 慧一郎 他 , 自己組織化マップを利用する分類済み階層の自動設定 , 電子情報通信学会 第
18回データ工学ワークショップ論文集 , 日本 , 電子情報通信学会データ工学研究専門委員会 , 2
007年 6月 1日 , Vol.DEWS2007, No.B7-8 , pp.1-7.

(58)調査した分野(Int.Cl., DB名)

G 06 F 17 / 30

J S T P l u s (J D r e a m I I I)