

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-55616

(P2010-55616A)

(43) 公開日 平成22年3月11日(2010.3.11)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06Q 30/00 (2006.01)</b>	G06F 17/60 3 2 6	5 B 0 7 5
<b>G06F 17/30 (2006.01)</b>	G06F 17/60 3 1 6	
<b>G09F 19/00 (2006.01)</b>	G06F 17/30 2 1 0 A	
	G06F 17/30 3 5 0 C	
	G06F 17/30 1 7 0 Z	
審査請求 未請求 請求項の数 4 O L (全 21 頁) 最終頁に続く		

(21) 出願番号 特願2009-191892 (P2009-191892)  
 (22) 出願日 平成21年8月21日 (2009. 8. 21)  
 (31) 優先権主張番号 61/092, 741  
 (32) 優先日 平成20年8月28日 (2008. 8. 28)  
 (33) 優先権主張国 米国 (US)  
 (31) 優先権主張番号 12/436, 067  
 (32) 優先日 平成21年5月5日 (2009. 5. 5)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 504407000  
 パロ アルト リサーチ センター イン  
 コーポレイテッド  
 アメリカ合衆国 カリフォルニア州 9 4  
 3 0 4 パロ アルト カイオーテ ヒル  
 ロード 3 3 3 3  
 (74) 代理人 100082005  
 弁理士 熊倉 禎男  
 (74) 代理人 100067013  
 弁理士 大塚 文昭  
 (74) 代理人 100086771  
 弁理士 西島 孝喜  
 (74) 代理人 100109070  
 弁理士 須田 洋之

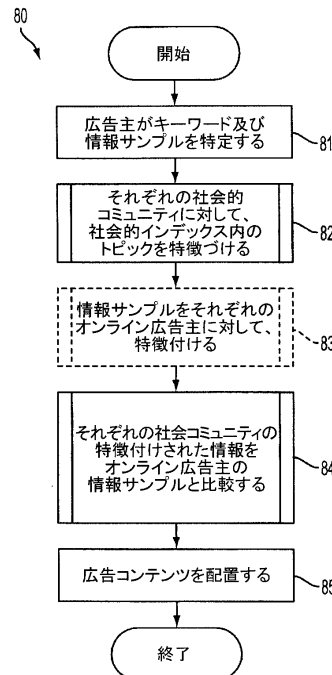
最終頁に続く

(54) 【発明の名称】 コミュニティに基づく広告語あいまいさ除去システム及び方法

(57) 【要約】

【課題】 広告語のあいまいさに関する問題を解決する。  
 【解決手段】 コミュニティに基づき広告語あいまいさ除去を行うための、コンピュータにより実行されるシステムであって、デジタル情報の前記記事と、各々が社会的コミュニティに関連させられた複数の社会的インデックスを含み、前記社会的インデックス各々が、それぞれ前記記事の一又は二以上の関連するトピックを含んでおり、プロセッサと該プロセッサによって実行されるコードを格納するメモリとを備えるコンピュータが備えられ、前記コンピュータは、前記社会的インデックスに基づいて広告内容に最も密接に適合する類似性を表示する前記社会的コミュニティを選択する社会的インデックス付与モジュールと、選択された前記社会的コミュニティの前記社会的インデックス内に備えられた前記トピックに関連した前記記事と共に前記広告内容を配置する広告処理モジュールと、を備えることを特徴とするシステム。

【選択図】 図9



## 【特許請求の範囲】

## 【請求項 1】

コミュニティに基づき広告語あいまいさ除去を行うための、コンピュータにより実行されるシステムであって、

デジタル情報の前記記事と、各々が社会的コミュニティに関連させられた複数の社会的インデックスとを含み、

前記社会的インデックスの各々が、それぞれ前記記事の一又はそれより多くに関連するトピックを含んでおり、

プロセッサと該プロセッサによって実行されるコードを格納するメモリとを備えるコンピュータが備えられ、

前記コンピュータは、前記社会的インデックスに基づいて広告内容に最も密接に適合する類似性を表示する前記社会的コミュニティを選択する社会的インデックス付与モジュールと、

選択された前記社会的コミュニティの前記社会的インデックス内に備えられた前記トピックに関連した前記記事と共に前記広告内容を配置する広告処理モジュールと、

を備えることを特徴とするシステム。

## 【請求項 2】

それぞれの社会的コミュニティに対して、広告内容を適合させる適合サブモジュールをさらに含み、前記適合サブモジュールは、

前記社会的コミュニティに関連する前記社会的インデックス内に含まれる前記トピックに関連する前記複数の記事を、前記社会的コミュニティの特徴によって、特徴付ける特徴判断サブモジュールと、

社会的コミュニティに対して、前記広告内容及び前記社会的コミュニティの特徴の類似性を評価する評価サブモジュールと、

を備える適合サブモジュールをさらに備えることを特徴とする請求項 1 に記載のシステム。

## 【請求項 3】

複数のオンライン広告主から広告内容の配置に対する複数の入札を受け取り、それぞれの前記オンライン広告主に関する前記広告内容を前記社会的コミュニティに適合させ、落札した入札に含まれる前記広告内容のみを配置する適合サブモジュールをさらに備えることを特徴とする請求項 1 に記載のシステム。

## 【請求項 4】

前記広告内容は、オンライン広告主に申し込まれる、少なくとも一つの製品及びサービスの情報サンプルの記述を含むことを特徴とする請求項 1 に記載のシステム。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本出願は、概して、オンライン広告に関し、特にコミュニティに基づく広告語あいまいさ除去システム及び方法に関する。

## 【背景技術】

## 【0002】

オークションに基づく検索広告が、オンライン広告の市場で優位を占めている。Yahoo! Inc. , Sunnyvale , CA によって主唱されたもののような、ダイレクト広告セールスは、ささやかな成功を収めただけである。このモデルでは、商業的バナー広告がウェブページに掲載され、ユーザのトラフィックをリンクされた広告主のウェブサイトに導くようにされていた。しかしながら、この方法は、不必要なコストがかかり、価格に関する透明性が欠けていた。

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0003】

10

20

30

40

50

広告オークションは、オンライン広告をリアルタイムに近い広告主間競合に開放することによって、価格についての障害を除去した。例えば、その開示が引用により組み入れられる、2001年9月4日に特許された米国特許第6,285,987号は、計算エージェントを用いて、広告からのキーワードをウェブサイトに対し符合させることによってスペースを獲得することを開示している。同様に、Yahoo!によるサービスであるOvertureは、最初にオンライン検索を広告と一体化するものであった。Overtureは又、検索クエリに広告を適合させるようにした取得メカニズムを用いている。それ以降、別の手法も開発されてきた。例えば、Google Inc., Mountain View, CAにより提案されたAdSenseは、オークションによる決定に応じて、広告をウェブページの異なる部分に配置する。オンライン広告に適合させるために用いられるキーワードを得ようとする広告主、若しくは広告が、ウェブページの内容に対して適合している一般的な検索クエリに関連させられる。

10

#### 【0004】

目標を持たない広告は、含まれる内容に関係なく、ウェブページ上に広告を配置するものであり、目標を持った広告は、密接な関係があるとみなされる広告のみを配置するものである。最近、クリック率の低下傾向に応じて、効果的な広告の目標付けが新しい役割をもつようになった。例えば、2007年11月12日発行のBusiness Weekの38ページに記載されたHolahan, C.による“広告が多すぎるため、クリックが僅かになる”によると、2006年には、Yahoo!, Microsoft及び AOLのような、主要なウェブポータルのある先に対する、目標を持たないバナー広告のクリック率は、0.75%から0.27%に落ち、バナー広告に対するウェブ全体での平均クリック率は、約0.2%に過ぎなかった。このクリック率の落ち込みによって、目標を持った広告に対する関心が高まってきた。

20

#### 【0005】

それにも関わらず、オンライン広告を効果的に目標付けすることは、困難なままである。図3から8を参照して以下で説明されるように、例えば、ルータのような本質的にあいまいな単語は、異なる市場を区別してオンライン広告を適切に目標付けするのに十分な内容を欠如している。そのために、従来手法は、組み合わせられるキーワードの全てをウェブページが含むようにし、ウェブページの各々に配置するためにキーワードの組み合わせを採用する。しかしながら、効果的な組み合わせを最初に特定し、次いで、効果を確認するためには、多数回のオークションによって評価をしなければならないので、言葉の組み合わせは、事業費用を増加させ、追加的な複雑性を生み出す。

30

#### 【0006】

効果的な目標付けは又、広告の市場価格決定、及び、活動性に影響を与える。合理的な広告主は、真正の製品及びサービスに結び付いたキーワードを取得しようとする。一方で、広く“スパマー”と呼ばれる非道徳的な広告主は、あまり知られていないキーワードを探す。このあまり知られていないキーワードは、取得が容易であり、実際に用いられるキーワードとは無関係な偽者の広告、若しくは、“スパム”広告の配置を可能にする。そのような広告は、十分な競争のない場所に存在する。反対に、オークションの価格が高過ぎる場合には、人工的に取得額が吊り上げられた価値あるキーワードが存在する市場において見られるように、オンライン広告は、一般に開放されず、見捨てられてしまう。経験のない取得希望者は、金銭上の損失によって排除されるまで、価格の吊り上げを継続してしまふ。その結果として、経験豊富な広告主が、特定ユーザのコミュニティに参与するか、若しくは、これら広告主のキーワードを調整することによって、市場を分割するようになる。

40

#### 【0007】

最後に、一つのウェブページでは、目標付けのための唯一の基準とするには、間違いを生じ、正確性が不十分であることがわかる。目標付けの論理が、特定のウェブページの内容のみに基づいている場合には、広告のターゲティングをするキーワードに対し誤った肯定判断をし、不適切若しくは役立たない広告の場所に導くことになる、という問題が

50

発生する。従来の手法は、キーワードを組み合わせることによって、誤った肯定判断をする率を下げ、極端に広い範囲にわたるトピックを持つコミュニティに關与することを避けており、これは、広告機会を逃す危険を伴う。したがって、効果的な目標付けは、ウェブページの内容以上に熟慮する必要があり、従来の方法では、依然として不十分である。

【0008】

例えば、その開示が引用により組み入れられる、2001年7月31日に特許された米国特許第6,269,361号は、検索エンジンによって生成された検索結果リスト上の位置に影響を与えるためのシステム及び方法を開示している。ウェブサイトのプロモータは、検索結果リスト作成のための検索リスト記載を定義し、プロモータのウェブサイトに関連する検索語を選択し、インターネット検索エンジン上における検索リスト記載の位置に影響を与えるようにすることができる。代替的な検索語を提案してもよい。続いて、ユーザが、検索語を入力するときには、検索エンジンは、取得希望額若しくはランク値のようなパラメータによって影響を受けた配置で、プロモータのリスト記載を用いて、検索結果のリストを生成することができる。しかしながら、検索語は、検索結果のリストによって提示され、不適切な場所の影響を受けた特定のウェブの内容に、強く結び付くことが予想される。

10

【0009】

その開示が引用により組み入れられる、2004年3月25日に公開された米国特許公開番号第2004/0059708号は、関連性のある広告を行う方法及び装置を開示する。目標文書の内容を分析することによって、広告に関する目標付け情報を特定し、トピックのリストを特定する。目標付け情報をトピックのリストと比較し、広告がターゲットドキュメントに関連するか否かを判断する。しかしながら、取得希望がトピックによって受け入れられるとしても、このトピックは、典型的には、他の誰かによって定義され、エラーを持つことがあり、しばしば不正確である。同様に、本質的にあいまいな言葉があると、悪い配置になることがある。

20

【0010】

その開示が引用により組み入れられる、2007年11月19日に公開された米国特許公開番号第2007/0260508号は、階層化した構造によって情報をリスト化するように広告を編成することを開示している。価格と価格決定のルールは、階層内にあるノードに割り当てられる。希望価格が、ノードのレベルに応じて提示され、ウェブコンテンツ及び広告は、階層的な構造の中で提供される。階層的な構造及びモードのラベルが通知なく変更されている場合であっても、広告主は、申し込み時に、ノード・階層アプローチに切り替えなければならず、従って、申し込みを意味のないものにする。用語の持つあいまいさの問題は、解決されない。

30

【0011】

また、ユーザの場所、訪問したウェブページの内容、及び検索履歴のような、ユーザのプロフィール情報が、目標付けを向上させるために用いられてきた。ユーザの情報を、目標付けプロフィールとして、永続的に格納することが可能である。代わりに、ユーザの興味を紹介するために、複数回ウェブサイトを訪問したという情報を集めることが可能である。しかし、プライバシーの提唱者が、連邦規定に申し立てを開始することによって、近年、強調されるように、プロフィール情報の利用は、違反的な侵害行為であると考えられることがある。

40

【0012】

例えば、その開示が引用により組み入れられる、2001年9月4日に公開された米国特許番号第6,285,987号は、インターネット広告システムを開示している。中央サーバが、どの広告を特定の閲覧者に表示すべきかを決定するために役立つウェブサイトの特徴、及び、その他の情報のような閲覧者に関する情報及び広告を格納する。前記その他の情報には、購買層の情報、及び様々な時間において閲覧者がアクセスしたサイトの内容に関する情報を含む。広告主の入札は、ユーザのプロフィールの特徴に基づいて、リアルタイムで評価される。

50

## 【課題を解決するための手段】

## 【0013】

一実施形態は、コミュニティに基づく広告語のあいまいさ除去のための、コンピュータによって実施されるシステム及び方法を提供する。社会的なコミュニティの各々に関連するデジタル情報の記事及び複数の社会的インデックスが維持される。それぞれの社会的インデックスは、各々又はそれより多い記事に関連するトピックを含む。広告する内容に最も密接に適合した類似性を示す社会的コミュニティを、社会的インデックスに基づいて選択する。選択された社会的コミュニティの社会的インデックスに含まれるトピックに関連する記事を持つ広告内容が、配置される。

## 【0014】

別の実施形態は、後述の詳細な記載からこの分野における当業者には、容易に明らかになるであろう。この詳細な記載には、意図されたベストモードを説明するための実施形態が記載されている。ここで理解されるように、本発明の思想及び範囲から外れることなく、別の実施形態及び異なる実施形態を実施すること、及び、実施形態の多くの詳細が、様々な明らかな観点における変形形態を実施することを可能にする。従って、図面及び詳細な説明は、例示的であり、発明を制限するものとみなされるべきではない。

## 【図面の簡単な説明】

## 【0015】

【図1】デジタル情報の意味付け及び情報検索のための典型的な環境を示すブロック図である。

【図2】図1の環境内で用いられる主要な要素を示す機能ブロック図である。

【図3】ルータに関する二つの異なるサンプルウェブページのスクリーンショットを例として示す図である。

【図4】二つの異なるルータのトピックに関する記事に対する特徴的な語及びnグラムを、例として示す図である。

【図5】二つの異なるルータのトピックに関する広告を例として示す図である。

【図6】図5のルータの広告に関するランディングページに関する最高位である特徴的な単語を、例として、示す図である。

【図7】ルータのトピックに関する共通の特徴的な語を、例として示す図である。

【図8】ルータのトピックに関する共通の特徴的な語を、例として示す図である。

【図9】一実施形態に合致するコミュニティに基づいて広告する語のあいまいさを除去する方法を示すフロー図である。

【図10】図9の方法を用いて、情報を特徴付けるルーティーンを示すフロー図である。

【図11】図10のルーティーンを用いて、粗いトピックモデルを作るルーティーンを示すフロー図である。

【図12】図11のルーティーンを用いて、特徴的な単語のスコアを任意に調整するルーティーンを示すフロー図である。

【図13】図9のルーティーンを用いて、社会的コミュニティを広告と比較するルーティーンを示すフロー図である。

【図14】図13のルーティーンを用いて、別の評価指標を判断するルーティーンを示すフロー図である。

## 【発明を実施するための形態】

## 【0016】

デジタル情報の意味付けと検索は、関連しているが、異なる動作である。前者は、デジタル情報のインフラストラクチャによって達成される意味付けに関連するものであり、インターネット、単独のコンピュータシステム、及び、開放したデジタル情報のリポジトリのような、公共のデータネットワークが含まれる。後者は、デジタル情報のインフラストラクチャからの情報の検索及び発掘に関連するものであり、このインフラストラクチャは、社会的なインデックスを通して、若しくは、別の情報源によって、トピックごとに統合される。図1は、デジタル情報の意味付け及び情報検索のための典型的な環境10を示す

10

20

30

40

50

ブロック図である。社会的なインデックスをつけるシステム 1 1 及びトピック検索システム 1 2 が、連携して機能し、意味付け及び検索をサポートする。

【 0 0 1 7 】

一般的には、デジタル情報は、デジタル形式で利用可能な情報の集積物である。情報の量に制約はなく、そのことは集積及びトピックの範囲が継続的に成長し、大きさやテーマに関する固定した境界が無いことを意味する。インターネットのようなデジタルデータの通信ネットワーク 1 6 は、デジタル情報を交換するためのインフラストラクチャを提供する。他のデジタル情報交換インフラストラクチャ、例えば私企業のネットワークであってもよい。

【 0 0 1 8 】

ネットワーク 1 6 は、情報集積物に存在する多様で分散的な情報源と、該情報集積物にアクセスするコンシューマとの相互接続を可能にする。著者、编者、協力者、外部寄与者は、継続的に、記事、ウェブページなどをネットワーク 1 6 に掲載し、これらは、ウェブサーバ 1 4 a、ニュース集積サーバ 1 4 b、投票可能なニュースサーバ 1 4 c、その他の情報ソースを通して、分散されたデータの集積物として維持される。これらのソースは、それぞれウェブコンテンツ 1 5 a、ニュースコンテンツ 1 5 b、及び、一般によって投票され若しくは " 審査 " されたコンテンツ 1 5 c を、ネットワーク 1 6 にアクセスするユーザに提供する。このアクセスは、例えばパーソナルコンピュータやその他のサーバであるユーザデバイス 1 3 a - c を介して行われる。ここでは分かり易くするために、ユーザデバイスについてだけ述べるが、サーバその他の非ユーザデバイス情報コンシューマも、同様に情報集積物内に維持された情報を調査、検索、使用することができる。

【 0 0 1 9 】

一般に、ユーザデバイス 1 3 a - c は、ウェブブラウザ若しくは類似のアプリケーションを実行する、ウェブ動作可能なデバイスであり、サーバ 1 4 a - c に対するインターフェースをサポートするとともに、サーバ 1 4 a - c との間における情報の交換及び検索をサポートする。ユーザデバイス及びサーバは、汎用目的のプログラム可能な計算デバイスに見られる中央処理装置、メモリ、入力/出力ポート、ネットワークインターフェース、及び、不揮発性格納部等を含む。ただし、これら以外のものを含んでもよい。さらに、サーバ 1 4 a - c の代わりに、若しくは、サーバ 1 4 a - c に加えて、別の情報ソースを含んでもよいし、ユーザデバイス 1 3 a - c の代わりに、若しくは、ユーザデバイス 1 3 a - c に加えて、別の情報コンシューマを含んでもよい。

【 0 0 2 0 】

デジタル情報の検索は、意味付けを補完する。一実施形態においては、トピックの検索システム 1 1 は、社会的インデックス付与システム 1 2 に統合される。社会的インデックス付与システム 1 1 によって提供されるトピックの編成は、トピック検索システム 1 2 によって有用に利用される。ただし、別のインデックス付けのソースを用いることもできる。例えば、その開示が引用により組み入れられる、2009年1月15日出願の、審査係属中である、一連番号 1 2 / 3 5 4 , 6 8 1 の同一出願人による米国特許出願に記載されたように、ユーザデバイス 1 3 a - c からの検索クエリが、インデックスの全体、あるいは焦点を合わせた単一のインデックスに対して実行され、そして、動的に焦点が合わせられ、トピック的に関連付けられた一連のインデックス及びそれらの先頭のトピック、若しくは、焦点が合わされた単一のインデックス内にある先頭のインデックスが、それぞれ、トピック検索システム 1 2 によって、サーチ結果を表示するために生成される。さらに、社会的インデックス付けのトピックに応じて構築された側面を利用して、検索された情報を持つ広告にたどり着くために用いられる目標付けを明確にすることによって、オンライン広告をトピック検索及びその他の検索動作に組み入れることができる。

【 0 0 2 1 】

ユーザの観点では、デジタル情報の検索のための環境 1 0 は、単一の情報ポータルのように見えるが、実際の裏側には、論理的には別々の、しかしながら統合された一連のサービスがある。オンライン広告は、アドオン方式で、検索に加えられる。図 2 は、環境 1 0

10

20

30

40

50

で用いられる主要な要素 20 を示す機能ブロック図である。この要素は、オンライン広告に焦点を合わせたものである。発見、予想、オリエンテーションなどの、別の関連した動作を行うために、追加の要素、若しくは、機能モジュールが必要となる場合がある。

#### 【0022】

要素 20 は、同じ計算プラットフォーム、若しくは、別々の計算プラットフォームによって実行される情報収集 21、広告 23、ユーザーサービス 26 モジュールに大まかに分類される。情報収集モジュール 21 は、入来コンテンツ 27 を開放された情報ソースから入手する。入来コンテンツ 27 は、新しいデジタル情報を情報集積物から継続的に取り入れるメディアコレクタによって集められる。入来コンテンツ 27 を、典型的には、構築されたリポジトリに格納するか、若しくは、実際のコピーを保持する代わりに、入来コンテンツに対するハイパーリンク若しくは引用を保存することによって、間接的に格納することができる。その上、入来コンテンツ 27 は、初めに格納されていた表現とは異なる複数の表現を含むことができる。異なる表現を用いることによって、タイトルを表示すること、記事の要約を表現すること、トピックの分類の軌跡を維持すること、及び、きめ細かいトピックモデルを導いて利用することなどが容易になる。記事中の言葉をせき止め、トークン形式化、マイナスの句点付け、大文字化、などをされて保存することができる。その上、社会的インデックス付けシステム 11 によって作成されるきめ細かいトピックのモデルは、入来コンテンツ 27 をかなり抽象化した変形版として表され、そこでは多くの単語が捨象され、単語の頻度が保持される。

10

#### 【0023】

入来コンテンツ 27 は、好ましくは少なくとも一つのトピックの社会的インデックス 29 の下で社会的インデックス付けを用いて、編集される。トピックの社会的インデックス 29 は、情報集積物内の全ての情報を対象とする、より大きい一連のトピックインデックス 33 の一部とすることができる。別の実施形態においては、その開示が引用により組み入れられる、同一出願人による、審査継続中の、2008年8月12日出願の一連番号 12/190,550 の米国特許出願のように、トピックのインデックス 29 は、社会的インデックスシステムを介して作られるエバグリーンインデックスとすることができる。エバグリーンインデックスは、新しい入来コンテンツ 27 が、一又はそれより大きいインデックスにおけるトピックの下に新しい情報が落ちるか否かをテストするために用いられる有限状態パターンのような、きめ細かいトピックモデルを含む。社会的インデックスシステムは、トレーニング教材を起動する管理された装置学習を、トピックインデックス 29 にそれぞれのトピック及びサブトピックに対するきめ細かいトピックモデルに適用する。トレーニングされた時点で、エバグリーンインデックスを、インデックスの外挿法のために用いて自動的に新しい情報を予め選択された主題領域のトピックの下で分類することができる。

20

30

#### 【0024】

図 9 に関連して以下に説明が始められるように、広告モジュール 23 は、オンライン広告の目標付けのあいまいさを除去する。広告モジュール 23 は、社会的なコミュニティ各々に対する一連のコミュニティの特性 32 を生成する特性付与サブモジュール 24 を含む。別の実施形態においては、特性付与サブモジュール 24 は、代わりに、それぞれのオンライン広告主に提供された広告 34 に対する一連の広告の特徴 35 を生成する。さらに、広告モジュール 23 は、類似性評価指標を判断することによって、広告コンテンツ 34 を比較して、適合させる比較サブモジュール 25 を含む。この類似性評価指標は、それぞれの広告主の製品若しくはサービスに最も強く関連しているトピックを持つ社会的コミュニティを特定する。別の実施形態においては、その開示が引用により組み入れられる、同一出願人によって 2009年5月5日出願され、審査継続中の米国特許出願一連番号 12/436,060 に記載されるように、オンライン広告は又、ユーザが特に関心を持つ記事若しくは社会的インデックスの一部に関連するトピックを特定することによって、範囲を広げられる。一連のトピックのインデックス 22、コミュニティ 32、広告 34、及び、広告の特徴 35 は、集約されたストレージ 28 に保持される。

40

50

## 【 0 0 2 5 】

最後に、ユーザサービスモジュール 2 6 は、フロントエンドをユーザ 3 0 a - b に提供し、一連のトピックのインデックス 2 2 及び入来コンテンツ 2 7 にアクセスし、一連のトピックのインデックス 2 2、若しくは、一つのトピックのインデックス 2 9 に対して、検索クエリを実行し、検索結果にアクセスする。さらに別の実施形態においては、それぞれのトピックのインデックス 2 9 は、“増大した”コミュニティとして知られるユーザのコミュニティに結び付く。そのコミュニティは、中心的な課題領域に対して継続的な関心を持つ。そのコミュニティは、それぞれのトピックに分類された記事に投票する 3 0 ことによって引用される情報を“検査する”。

## 【 0 0 2 6 】

オンライン広告のライフサイクルには、広告収入を稼ぐことに関連した三つのイベントがある。一つ目は、ウェブページに広告を表示するときに常に発生する表示イベントである。表示若しくは表現の回数は、CPM（千回の表示若しくは“千人”に対する表示当たりのコスト）プライスモデルの応じた計算の基礎になる。二つ目は、ユーザが表示された広告の上をクリックするときに常に発生するクリックを介するイベントである。クリックを介するイベントの数は、通常の場合には、CPC（クリック当たりのコスト）プライスモデルに応じた収入を計算するための基礎となる。三つ目には、記録若しくは商品の購入のような、ユーザが、広告主のサイトに対して行動を取るときに常に発生する会話イベントである。一般的には、広告主は、会話イベントに関する統計を管理しており、期待される会話の比率及び潜在的な収入の見積もりに基づいて、CPM、若しくは、CPC 広告に対して支払う額を見積もる。

## 【 0 0 2 7 】

適切に適合したコンテンツに対する目標付け、及び、適合したウェブページ内でのオンライン広告の適切な配置は、潜在的な収入に影響を与える。収入を得る機会が、効果のない目標付けによって失われることがある。配置の悪い広告は、目立たず、見落とされるか、単に無視されることがある。社会的インデックスは、目標付けを改良し、オンライン広告の配置をよりよいものにすることができる。社会的インデックス付けは、トピックに応じて編成された主題の領域の記事を表示し、ユーザが、すばやく特定のトピックに関する情報にアクセスするために役立つ。その上、一人のユーザより多い観衆が興味を持つ記事を編成するコミュニティによって、社会的インデックスは、用いられる。従って、社会的インデックスは、オンライン広告に用いることができ、ユーザをコミュニティに分類する情報サービスに広告をターゲットする。このことによって、潜在的な広告収入を増加させる一方で、ユーザを特定する行為を避けることになる。

## 【 0 0 2 8 】

従来は、広告主に提供されたキーワードのみによって、オンライン広告を目標付けしていたので、文意的なあいまいさ及び不適切な配置のリスクを招いていた。図 3 は、ルータに関する二つの異なるサンプルウェブページ 4 1、4 2 のスクリーンショット 4 0 を例として示す図である。例えば、キーワード“ルータ”は多くの意味を持つ。ある意味においては、ネットワークルータに関する Wikipedia のウェブページ 4 1 の説明によると、ルータは、ネットワークメッセージの交通を管理するデバイスである。別の意味においては、木工機械に関する特殊なルータツールに関するウェブページ 4 2 の説明によると、ルータは、木工のパワーツールである。

## 【 0 0 2 9 】

ルータに関する記事の間で、これらの二つのウェブページ 4 1、4 2 は、ユーザの異なるコミュニティにとっての潜在的な関心事になる。ネットワークルータのウェブページ 4 1 は、コンピュータ通信、及び、ネットワーク通信を構築するためのルータの使い方に興味を持つ人々の関心事になる。木工のルータのウェブページ 4 2 は、木工ツールに興味を持つ建具職人の関心事になる。図 4 は、二つの異なるルータトピック 5 1、5 2 に関する記事に対する特徴的な単語 5 3、5 4 及び最高位の n グラム 5 5、5 6 を、例として示す図である。木工ルータ 5 1 を表す表は、ルータに興味を持つ建具のコミュニティに関する

10

20

30

40

50



いくつかの記事にわたって、判断された。特徴的な単語の表 5 5 は、記事内に現れる全ての単語を特定することによって、判断された。このシステムは、単語出現頻度 - 逆文書出現頻度 (TF-IDF) の重み付けを用いて、ルータに関する記事内に存在するそれぞれの単語の出現頻度を数えた。選択した記事内にあるそれぞれの単語の出現頻度を、より大きな情報集積物内にある言葉の頻度で割って、独自性を見積もりを算出する。

#### 【0030】

この計算は、より大きな情報集積物内よりも、一連の記事内の方が、頻度が大きくなる単語を特定する、つまり、主題において、特徴的な単語を特定する。この数字は、100で規格化して、表に分類した。木工ルータに関する記事内の単語の中で、最も特徴的な単語は、“ツール”であり、これに続いて、“ルータ”“アクセサリ”“デフォルト”“マキタ”などから“ショップ”までがあった。最も頻出な n グラム 5 5 の表は、同じ一群の記事の中で見つかった隣接語の列を用いて、特徴的な単語で構成される。最も頻出の n グラムは、“パワーツール”続いて“カットアウトツール”及び“ポータアクセサリ”などであった。ネットワークルータを表す表は、ルータに関心を持つコンピュータネットワークのコミュニティに関連する幾つかの記事にわたって、同様に判断された。同じ計算が実行されて、その結果は、特徴的な単語 5 4 及び最高位の n グラム 5 6 の得られたテーブル内において、顕著な違いを示す。最も特徴的な単語は、“データシート”続いて“ジュニパー”“ネットワーク”“シリーズ”“サービス”などから“シスコ”までであった。最高位の n グラムは、“ジュニパーネットワーク”続いて、“ファイアウォール i p s e c”“セキュリティプロダクト比較”などであった。これらのリストは、コミュニティの関心事を示す、情報集積物内にある記事に由来している。計算した数字は、情報集積物の大きさが大きくなるにつれて、安定する。サンプリング技術をドキュメント群にわたって用いることができ、計算時間を制限するだけでなく、例えば、サブピックによって、より新しい記事に対して、若しくは、高くランク付けをされた記事に対して、故意に、計算にバイアスをかけることもできる。

#### 【0031】

これらの表 5 3 - 5 6 は、一連の記事の主題を特徴付けるように意図された表現の例である。類似の表現は、“タグの雲”のような視覚的な表現の生成の基礎になる。関連する一群の計算は、特定の言葉の“近隣”に現れる言葉を探す。例えば、情報の手がかりとしての役目を果たす記事内、若しくは、ドキュメント群にわたって、“ルータ”近隣の 4 0 語のウィンドウに現れている言葉の頻度を反映するように、探す。さらに別の表現を用いてもよい。

#### 【0032】

文意のあいまいさの解消方法は、例を用いて説明することができる。ネットワークリキデータは、ウェブ上で営業を行う、ルータのような改造機器を提供する架空の企業である。ルータヘッドクォータは又、木工のルータのアクセサリ及びツールに特化する架空の小売店である。両方の会社は、キーワード“ルータ”を用いて、2008年1月にウェブ上で広告された。図 5 は、これらの会社に対する二つの異なるルータのトピックに関する広告 6 1、6 3 を、例として示す図である。図 5 は又、それぞれの広告の上でクリックすることによって、たどり着く“ランディングページ”と呼ばれるウェブページ 6 2、6 4 を示す。ルータヘッドクォータのランディングページ 6 2 は、約 2,800 語を持ち、ネットワークリキデータのランディングページ 6 4 は、800 語を持つ。これらの会社の広告する製品は、全く異なっているので、製品は、異なる関心を持つ人々にアピールするであろう。

#### 【0033】

市場の区分は、広告をユーザのコミュニティに適合することを要求する。最高位の特徴的な単語のリストは、市場の区分に関する情報源の一つであり、図 3 に関連して上述したように、判断される。図 6 は、図 5 のルータの広告に関するランディングページ 6 2、6 4 に関する最高位である特徴的な単語を、例として、示す図である。ルータヘッドクォータに関する最高位の特徴的な単語を示す表 7 1 は、左側にあり、ネットワークリキデ

10

20

30

40

50

タに関する最高位の特徴的な単語を示す表 7 2 は、右側にある。図 7 及び 8 は、ルータのトピックに関する共通の特徴的な単語を、例として、示す図である。図 7 を参照して示される表を最初に参照すると、ネットワークリキデータに関する広告は、ネットワーキングのコミュニティとの共通点を多く持ち、ルータヘッドクウォータに関する広告は、木工のコミュニティとの共通点を多く持つ。

#### 【 0 0 3 4 】

この評価指標の変形は、より表示的であってもよく、幾分簡単に計算を行う。一般的に、解析に利用できる情報集積物は、製品に対してよりもコミュニティに対しての方が、より大きくなると予期される。したがって、特徴的な単語の計算は、コミュニティに対しての方が、広告に対してよりも意味があるであろう。図 8 に関連して示される図を次に参照すると、広告のランディングページに現れるユーザコミュニティに関する特徴的な単語が、判断される。この場合には、ネットワークコミュニティからの 8 つの特徴的な単語が、ネットワークリキデータのランディングページ 6 4 に現れ、木工のコミュニティの最高位の特徴的な 2 5 の単語全てが、ルータヘッドクウォータのランディングページ 6 2 に現れた。

10

#### 【 0 0 3 5 】

広告は、トピック検索の結果と共に表示するか、別の動作と協同して表示してもよい。例えば、ユーザが、ニュースページ、若しくは、ウェブページにあるインデックス内のトピック上に存在する、ストーリー若しくはトピックをたどることによって、情報を減らすことができる。これらの両出発点は、より深い位置に編成された社会的インデックスの情報にユーザを導き、トピックによって情報を表示するいずれのページも、潜在的に広告に役立つ場所である。一般的に、ユーザが、より深い位置にある社会的インデックスに行くほど、専門化したトピックが増え、潜在的に、より正確に目標付けをした広告になる。このような情報への道は、直接的に記事を調べることを介して、トピックの検索を実行することによるか、若しくは、情報を探して、情報にアクセスする別の手段によって、結果的に、情報的な軌跡若しくはトピックの軌跡をたどる。

20

#### 【 0 0 3 6 】

したがって、共通のトピックの社会的インデックスからのコミュニティ情報を用いて、オンライン広告がオンラインコミュニティに目標付けされる。目的は、言葉の意味及び感覚におけるあいまいさによる影響を受ける、広告の不適切な配置を減らすことである。社会的インデックス付けの文意において、コミュニティは、主題領域の周りで編成される。それぞれのコミュニティは、システムが周期的に収集して、解析する一連のオンラインの情報源を特定する。社会的インデックスは、一連のインデックスの入力を具体化する。その入力は、収集された記事に自動的に適合され、それが、トピックによって素材の表現を編成できるようにする。システムに投票するコミュニティは、収集された記事を質によってランク付けするために用いられる。

30

#### 【 0 0 3 7 】

それぞれの社会的インデックスが、オンラインユーザの社会的コミュニティによって、選択されるトピックに応じて分類されたトピック的に関連した記事にわたって、コンテンツの解析を集める。したがって、オンライン広告は、特定のウェブページの内容若しくはユーザのプロフィール単独でターゲットされるのではなく、むしろ、提供される基礎的な製品若しくはサービスに対する関心を最も持とうとする社会的なコミュニティに適合される。図 9 は、一実施形態に合致するコミュニティに基づいて広告する語のあいまいさを除去する方法 8 0 を示すフロー図である。方法 8 0 は、例えば、一般的な目的でプログラムされた、サーバのようなコンピュータによって、一連のプロセス若しくは方法のステップとして、実行される。

40

#### 【 0 0 3 8 】

社会的コミュニティの情報を用いて、目標付けにおける誤った肯定判断を減少させることによって、広告収入の潜在性を改善することが可能である。全てのオンライン広告主は、広告オークションを成功的に落札するか、若しくは、別の配置基準を満たすことによ

50

て、配置される広告コンテンツ34を特定する。目標付けを促進するために、それぞれの広告主は、キーワード及び情報のサンプルを特定する(ステップ81)。広告の“ランディングページ”すなわち、クリックイベントに続いて提供されるウェブページ上に見つかるように、情報のサンプルは、提供される製品若しくはサービスの説明を含んでもよい。追加的な製品情報が又、広告主によって、提供される。

#### 【0039】

効率的なマーケティングのセグメント化をするためには、オンライン広告を関連ある社会的なコミュニティに適合させる必要がある。オンライン広告主に提供される情報のサンプル、及び、それぞれの社会的コミュニティに属する社会的インデックスの下に現れるトピックの両方を特徴付けることによって、その適合化を達成することができる。広告主各々の情報のサンプルを特徴付けること(ステップ83)によって、それぞれの潜在的な広告を、広告主の製品若しくはサービスと共通点を最もたくさん持つ社会的コミュニティに適合するようにする。潜在的な広告に適合する代わりにそれぞれの社会的コミュニティのトピックを特徴づけること(ステップ82)によって、同等の適合結果を得るが、計算の効率性を向上させることも魅力的である。社会的なコミュニティは、社会的インデックスを用いて、主題領域の周辺に編成される。それぞれのコミュニティは、一連のオンライン情報源を特定する。その情報源から、社会的インデックス付けシステム11は、周期的に新しい記事を集めて、解析する。それぞれの社会的インデックスは、トピックによって記事の表現を自己編成するために、集められた記事に適合する一連のインデックスの入力を具体化する。追加的に、コミュニティの投票は、質によって、集められた記事にランク付けをする。

10

20

#### 【0040】

一般的に、社会的インデックス付けを介して解析するために利用できる情報集積物は、社会的なコミュニティに関連するものとして、広告されるべき製品若しくはサービスに関する情報サンプルよりも、とても大きい。2009年1月15日に出願した一連番号12/354,681の同一出願人による米国特許出願に記載されたように、インデックス内にある情報の特徴付けを、例えばトピックに基づいた検索をサポートするような複数の目的で用いることができる。当該出願は、審査中であり、その公開公報は、参照により、本明細書に組み込まれているものとする。図10に関連して以下から記載するように、結果的に、一実施形態においては、それぞれの社会的なコミュニティに対する社会的インデックス内にあるトピックを特徴付けるが(ステップ82)、別の実施形態においては、それぞれのオンライン広告主に対する情報のサンプルを特徴づける(ステップ83)。

30

#### 【0041】

したがって、図13を参照して、以下に記載されるように、それぞれの社会的コミュニティの特徴付けされた情報は、オンライン広告主の情報サンプルと比較され(ステップ84)、そのサンプルから類似の評価指標が生成される。別の実施形態においては、広告主の特徴付けされた情報は、それぞれの社会的コミュニティに対する社会的インデックス内にあるトピックと比較される。それぞれの類似性評価指標を、直接的にオンライン広告主に、並びに、オンライン広告の場所を監視することによって、課金される出版者、広告プロモーター、若しくは第三者集団に、提供することができる。また、類似性評価指標を、アプリケーションプログラミングインターフェースを介して、第三者集団の広告サーバ、若しくは、別の外部システムに提供することができる。

40

#### 【0042】

社会的コミュニティを基にした目標付け広告の配置は、広告が適切で明確な文意で存在することを確保する。従って、広告コンテンツは、最も密接に適合する社会的コミュニティの社会的インデックスに関するトピックの下で存在する記事の中に配置される(ステップ85)。落札するキーワードの入札に基づいて、オークションスタイルの形式若しくは別の広告主間でのコンペティションを用いて、或いは、別の配置基準によって、広告を配置することができる。別の実施形態においては、広告主のキーワードは、社会的なコミュニティの情報にキーワードオークションを用いなくて、適合される。例えば、広告主は、

50

広告及び情報サンプルを提供し、広告の配置に入札する。広告主は、関連するコミュニティに適合し、広告の選択がオークションのみによって判断されるであろう。クリックの見込み及び関連する技術は、キーワードオークションと同様の方法で提供されるであろう。

【0043】

別の実施形態においては、オンライン広告は、明確な社会的コミュニティよりもあいまいなコミュニティを用いて、目標付けされる。あいまいなコミュニティは、例えば、クリック、及び、場合によっては会話イベントについての統計を取ることによって、秘密裏に判断され得る。クリック率が高いウェブページの特徴を解析し、特徴的な単語、若しくは、別の兆候によってウェブページを集める。あいまいなコミュニティに対する関心を持つであろうトピック若しくは主題は、このようにして、特定される。

10

【0044】

オンライン広告によって提供されるサンプル、若しくは、それぞれの社会的コミュニティに社会的インデックスの下で現れるトピックのいずれか、又は、それらの両方を、類似性評価指標の比較及び判断の前に特徴付けなければならない。図10は、図9の方法80を用いて、情報を特徴付けるルーティーン90を示すフロー図である。それぞれの情報源、つまり、広告主の情報のサンプル、若しくは、社会的コミュニティの社会的インデックスからのトピックのいずれか、からの情報が、一又はそれより多い特徴付け技術を用いて、特徴付けされる(ステップ91-95)。

【0045】

例えば、図11及び図12に関連して以下に記載されるように、一連の特徴的な言葉のトピックモデルが、情報の中で現れる全ての言葉の頻度特定することによって、一連の特徴的な単語のトピックモデルを形成する(ステップ92)。別の実施形態においては、一連の隣接する特徴的な単語である、最も頻繁に現れるnグラムを、一つずつの特徴的な単語の代わりに、決定することができる。別のサンプリング技術が、情報に対して用いられて、計算時間を制限するだけでなく、より新しい記事に対して、若しくは、高いランク付けをされた記事に対して、例えば、“タグクラウド”のような視覚的な表示を用いて、サブトピックに応じて、故意に計算にバイアスをかけることができる(ステップ93)。関連する一群の計算は、任意の言葉の“近隣”に現れる言葉を評価する。例えば、特定の言葉の40語のウィンドウに現れている言葉の頻度を決定して、“情報の手がかり”を提供する(ステップ94)。その上、前述の一連の特徴付け技術は、網羅的ではなく、別の技術を採用し、受け入れられる結果を提供することができる。

20

30

【0046】

特徴的な単語は、オンライン広告主に提供された情報のサンプルを特徴付ける手段になる以外にも、それぞれの社会的なコミュニティに属する社会的インデックスの下に現れるトピックを特徴付ける手段になる。特徴的な単語、若しくは、“粗い”トピックモデルは、より大きな情報情報集積物内で、別の単語よりも頻繁に現れる単語を特定する。したがって、トピックモデルは、主題領域において、特徴的な単語を見つける。

【0047】

粗いトピックモデルを、オンライン広告の目標付けの前に予め計算することができる。図11は、図10のルーティーン90と共に用いる、粗いトピックモデルを作るルーティーン100を示すフロー図である。粗いトピックモデルは、特徴的な単語、及び、それぞれの特徴的な単語の相対的重要性を反映したスコアを含む。

40

【0048】

特徴的な単語は、トピックに関するテキストを区別するのに役立つ、典型的には、適切な情報集積物内の記事から選択された言葉である。その情報集積物は、ウェブページ、電子書籍、若しくは、その他の印刷された素材のように利用できるデジタル情報を含むことができる。最初に一連の若しくはランダムに抽出した記事を、情報集積物から選択する(ステップ101)。特徴的な単語の基準及び発生頻度は、選択した記事から抽出される(ステップ102)。インデックス29内のトピックに対する基準は、インデックス29の情報集積物に応じて、判断される。一式のインデックス22に対する基準は、個別のイン

50

デックス29の全てに対するコーポラである、システム全体の情報集積物にわたって計算される。待ち時間を縮めるために、基準を満たすそれぞれの特徴的な単語の存在の頻度を予め計算することができる。一実施形態においては、インデックス内にあるトピック下に存在する記事の数が、例えば一時間ごとに計測される。周期的に、記事の数が、10パーセントのように予め決められた量の差で、変化した場合には、存在の頻度が再度決定される。

#### 【0049】

次に、一般的にユーザに選択された一連のポジティブトレーニングの例が得られる（ステップ103）。このポジティブトレーニングの例は、前述されたように、エバグリーンインデックスに関するきめ細かいトピックモデルを構築するときに管理された学習の際に用いられたものと同じ一連の記事とすることができる。別の実施形態においては、きめ細かいトピックモデルと合致する記事の例が、ポジティブトレーニング例の代わりに用いられる。特徴的な単語が、ポジティブトレーニング例から抽出されて、ポジティブトレーニング例の中にあるそれぞれの特徴的な単語の存在する頻度が、決定される（ステップ104）。計測若しくはスコアが、例えば、TF-IDFの重みを用いて、それぞれの特徴的な単語に割り当てられる。この語頻度-逆文献頻度（TF-IDF）の重みは、ベースライン内にあるそれぞれの特徴的な単語の出現頻度に対するポジティブトレーニング例の中にあるそれぞれの特徴的な単語の出現頻度との比を特定する（ステップ105）。図12に関連して以下に示すように、それぞれの特徴的な単語のスコアを調整して（ステップ106）、トピックに対して特徴的な単語の重要性を向上させるか、若しくは、低下させることができる。最終的には、オンライン広告要求処理段階で用いるための特徴的な単語及びスコアの表が生成される（ステップ107）。特徴的な単語及びスコアをリストにして、この表を整理し、集めることができる。別の型の表も可能である。

10

20

#### 【0050】

異なる“アクター”が、粗いモデルを作るという動作を実行する。システムは、一連の、若しくは、ランダムに抽出した記事を選択し、基準を判断する。人間であってもよいインデックス管理者は、ポジティブトレーニングの例を選択する。残りのステップが、システムによって実行される。ポジティブトレーニング例の選択を、前もって、ルーティーンにおけるあらゆる別のステップより先に完了することができる。ルーティーンは、入力として、情報集積物内にある一連の記事、及び、ポジティブトレーニング例として指定された情報集積物から一連の記事を受け取る。ポジティブトレーニング例が、きめ細かいモデルと合致する記事である場合においては、同じことが認識される。ここで、2009年1月27日に出願した一連番号12/360,825の同一出願人による米国特許出願に記載されたように、きめ細かいモデルは、トピックラベルに基づいたきめ細かいパターンを作る“初期トレーニングアルゴリズム”に由来する。当該出願は、審査中であり、その公開公報は、参照により、本明細書に組み込まれているものとする。きめ細かいトピックモデルを作るこれらの二つのアプローチは、“初期トピックトレーニング”及び“例に基づくトピックトレーニング”と呼ばれる。

30

#### 【0051】

それぞれの特徴的なスコアは、存在頻度の生の比率を反映している。図12は、図11のルーティーン100を用いる特徴的な単語のスコアを、任意に調整するルーティーン110を示すフロー図である。試行錯誤的に、それぞれの特徴的な単語のスコアを文脈に応じて多くの方法で調整することができる。例えば、めったにない言葉、つまり、情報集積物、若しくは、引用した一連の素材内で最小の回数より少ない回数で現れる言葉のスコアを、例えば、25パーセント圧縮、若しくは、縮小することができる（ステップ111）。同様に、文字数の最小となる閾値より短い言葉のスコアは、同様のパーセントで圧縮することができる（ステップ112）。なぜなら、短い言葉は、トピック的に重要な意味を持たない場合が多いからである。反対に、ラベル若しくはタイトル内にある言葉は、強くトピックの内容を反映しているので、それらのスコアは、例の中に存在する言葉に比べて何倍も押し上げられて、増加する（ステップ113とステップ114それぞれ）。典型的

40

50

には、全てのラベルの言葉は、特徴的な単語として含まれる。最後に、ラベル語に近接して存在する言葉、つまり、近接語及び、セットウインドウ内にあるラベル語の周囲に存在する“近隣”語のスコアは、押し上げられる（ステップ115）。規格化された閾値が、近接語及び近隣語の選択中に採用される。最大スコアの8パーセント及び15パーセントの初期の閾値がそれぞれ、近接語及び近隣語に、8語のセットウインドウの大きさを伴って、それぞれ採用される。別の代表的な閾値及び長さを用いることもできる。最終的に、特徴的な単語のスコアは、規格化される（ステップ116）。最高のスコアを持つ特徴的な単語は又、最も固有の言葉であり、そのスコアを100パーセントに設定する。残った特徴的な単語のスコアは、最高のスコアを基にして、拡大縮小される。したがって、特徴的な単語の選択を完了した時点で、インデックス内にあるそれぞれのトピックは、情報集積物からサンプリングされた題材に応じて、規格化された特徴的な単語に関連して表現された粗いトピックモデルを持つ。

10

**【0052】**

最もよく適合した候補となるコミュニティが、それぞれの社会的コミュニティの特徴的な情報を、オンライン広告の情報サンプルと比較することによって、選択される。相性の良し悪しは、比較動作を定量評価する類似性評価指標によって測定される。図13は、図9のルーティーン80を用いて、社会的コミュニティを広告と比較するルーティーン120を示すフロー図である。それぞれの社会的なコミュニティによって、特徴付けられた情報が、一又はそれより多い比較技術を用いて、比較される（ステップ121 - 129）。この比較技術は、以下のものを含む。

20

**【0053】**

情報サンプル内に現れるコミュニティの特徴的な単語の計数（ステップ122）。最高位の特徴的な単語は、上述のように、それぞれの社会的なコミュニティごとに計算される。広告主のランディングページ、若しくは、別の情報サンプル内に現れる最高位の特徴的な単語の数は、製品若しくはサービスに対する社会的コミュニティの潜在的な興味を反映する類似性評価指標になる。最も関連ある広告は、対応する情報サンプル内に最も多く現れる特徴的な単語を含む。

**【0054】**

コミュニティの特徴的な単語の数の重み付け（ステップ123）。追加の重みをそれぞれのコミュニティの最も特徴的な単語に与える。類似性評価指標は、広告主のランディングページ内に現れる重み付けした特長的な単語の数である。別の実施形態においては、追加的な重みは、ランディングページ、若しくは、別の広告主の情報サンプル内に繰り返し現れる特徴的な単語に与えられる。

30

**【0055】**

ベクトル空間距離（ステップ124）。ベクトル空間に対する確率的アプローチは又、関係性の評価に用いられる。単語ベクトルは、ベクトル空間内にある点として扱われる。二つの点がベクトル空間内で近いほど、関係性の評価が高くなる。関係性の評価は、語の使い方に基づいて、任意の大量のドキュメント内に、同じ特徴的なサンプルから任意の記事が引き出される確率として、再計算される。例えば、二つのベクトルの距離は、それぞれの語の係数の差の二乗の平方根を取ることによって計算される。単語ベクトルが、おおむね同じ係数を全ての語に対して持つならば、二つの語のベクトルは、近くにある。完全に一致したとき、距離は、ゼロであろう。別の実施形態においては、次元低下技術を用いて、高い相関がある語を組み合わせることによって、計算上の語の数を減らすことができる。

40

**【0056】**

ベクトル空間の写像（ステップ125）。語のベクトルの写像が、測定される。語のベクトルの写像が、類似するほど、関係性が増す。それぞれの語の重要性が、重みによって判断され、関係性の評価が、語のベクトルのドットプロダクトによって、判断される。ドットプロダクトは、広告主の製品若しくはサービスの情報を表現する語のベクトル、及び、コミュニティの関心を表現する語のベクトルの間で計算される。製品若しくはサービス

50

情報の最も高いドットプロダクトを持つコミュニティは、そのプロダクトに最も高い関心があるものとして評価される。

【 0 0 5 7 】

n グラム変動 (ステップ 1 2 6)。ベクトル空間の写像は、基本的な要素として個別の語を捕らえる。しかしながら、その意味を、言葉の並びによって表現されるフレーズを用いて、代わりに伝えることができる。一般的な n グラムを、ドキュメントのセット、若しくは、特徴的な単語の頻度及び特徴的な n グラムの頻度の組み合わせに対して、判断することができる。

【 0 0 5 8 】

インデックスパターンの適合に基づいた評価指標 (ステップ 1 2 7)。コミュニティのトピックのインデックスの構造、及び、前もって行われるトピックを認識するパターンの計算が、抽出される。それぞれのトピックに対するパターンを広告主の情報に対して適合するように、それぞれのコミュニティのインデックスは、トピックに応じて、広告主の情報のサンプルを分類する。また、様々な評価指標をそれぞれの社会的コミュニティに対して、計算する。例えば、ある類似性評価指標は、パターンがランディングページに適合するサブトピックの数を単純に数える。適合するパターンの数、若しくは、パーセンテージは、類似性評価指標になり、広告が、社会的コミュニティの関心に関連する場合において、増加する。

10

【 0 0 5 9 】

別の評価指標 (ステップ 1 2 8)。類似性評価指標は、社会的コミュニティと広告主の相性の良し悪しを、測定する。図 1 4 に関連して以下に記載されるように、別の評価指標が、類似性評価指標の組み合わせを用いて、生成される。

20

【 0 0 6 0 】

上述の一連の比較技術は、網羅的ではなく、別の技術を採用して、受け入れられる結果を提供することができる。類似性評価指標を、広告の配置を用いて、戻す (ステップ 1 3 0)。

【 0 0 6 1 】

追加的な評価指標を、別の情報源から取得し、社会的コミュニティの選択を手助けすることができる。図 1 4 は、図 1 3 のルーティーン 1 2 0 を用いて、別の評価指標を判断するルーティーン 1 4 0 を示すフロー図である。別の評価指標は、一又はそれより多い技術を用いて判断される (ステップ 1 4 1 - 1 4 6)。この評価指標は、以下のものを含む。

30

【 0 0 6 2 】

コミュニティの目的に関するテキスト (ステップ 1 4 1)。それぞれの社会的コミュニティは、コミュニティの目的を表現するテキストの記載を持つことができる。これらの記載は、潜在的に広告するコミュニティを選択するとき、広告主によって用いることができる。広告のキーワードを見つけるために、この記載を検索し、この記載を、広告のキーワードに適合する。また、この記載の中にある単語を広告主の製品若しくはサービスの記載の中にある単語と比較し、用語的な重なりを指標を提示する。

【 0 0 6 3 】

活動統計 (ステップ 1 4 2)。例えば、アクティブのメンバー、若しくは訪問者に応じて、コミュニティの大きさに関するデータを含む、社会的コミュニティのオンラインプロフィールから、統計を計算することができる。

40

【 0 0 6 4 】

人口動態 (ステップ 1 4 3)。広告主は、商品若しくは製品を、社会的コミュニティの人口動態の情報に適合させる。その人口動態の情報は、コミュニティの社会的インデックス内にあるトピックから得られるか、若しくは、アンケート及びサンプリングを集めて得られる。コミュニティの人口動態は、重複なく、実際のユーザから判断され、コミュニティのトピックに関心がある人を示す人口サンプルを調べることによって、個人のプライバシーが保護される。

【 0 0 6 5 】

50

特徴的な単語の適合（ステップ144）。評価指標は、全てのトピックにわたって、社会的コミュニティによって用いられる特徴的な単語を、広告主の情報の中にある語と比較することができる。例えば、評価指標は、広告主の情報のサンプルに適合する特徴的な単語の数を数えることができる。

【0066】

適合するトピックを数える（ステップ145）。社会的コミュニティのトピックのパターンに関連するパターンを、広告主の情報のサンプルに適合することができる。例えば、それぞれのコミュニティに対して、広告主の情報のサンプルに適合するトピックパターンの数若しくはパーセンテージを評価指標として、用いることができる。

【0067】

上述の一連の別の評価指標は、網羅的ではなく、別の評価指標を生成し、コミュニティ選択プロセスに含めることができる。それぞれの別の評価指標は、広告の配置を用いて、戻される（ステップ146）。

【0068】

本発明が、特にその実施形態を参照して示されて記載されているが、本分野の当業者は、形式及び詳細において、本発明の思想及び範囲を外れることなく、既に述べたあるいはその他の変更を行うことができることを理解するであろう。当該出願は、審査中であり、その公開公報は、参照により、本明細書に組み込まれているものとする。

【符号の説明】

【0069】

- 10 環境
- 11 システム
- 12 トピック検索システム
- 13 a - c ユーザデバイス
- 14 a ウェブサーバ
- 14 b ニュースの集合サーバ
- 14 c 投票可能なニュースサーバ
- 15 a ウェブコンテンツ
- 15 b ニュースコンテンツ
- 15 c “審査された”コンテンツ
- 16 ネットワーク
- 20 主要な要素
- 21 情報収集モジュール
- 22 一連のトピックのインデックス
- 23 広告モジュール
- 24 特性付与サブモジュール
- 25 比較サブモジュール
- 26 ユーザサービスモジュール
- 27 入来コンテンツ
- 28 ストレージ
- 29 インデックス
- 30 投票する（ユーザ）
- 32 コミュニティ
- 33 トピックインデックス
- 34 広告コンテンツ
- 35 特徴
- 40 スクリーンショット
- 41 ウェブページ
- 42 ウェブページ
- 51 ルータ

10

20

30

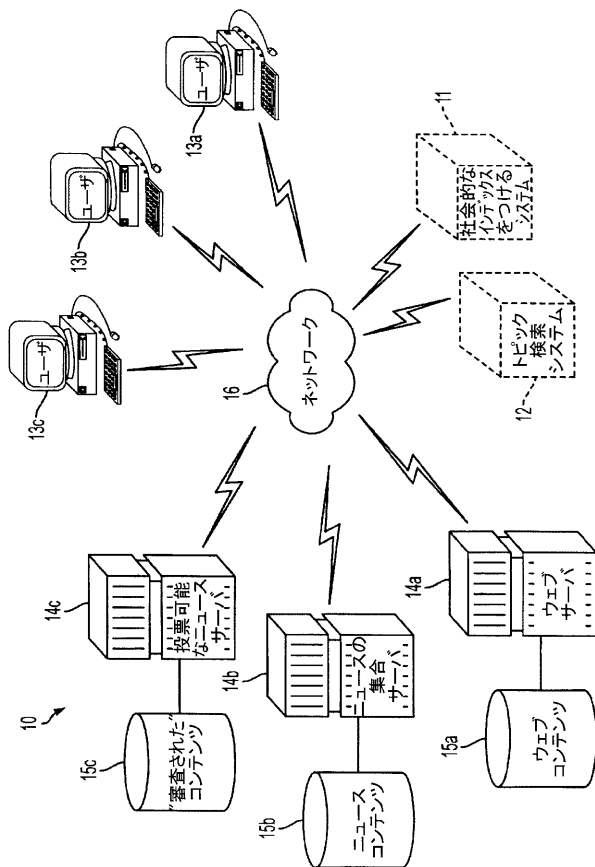
40

50

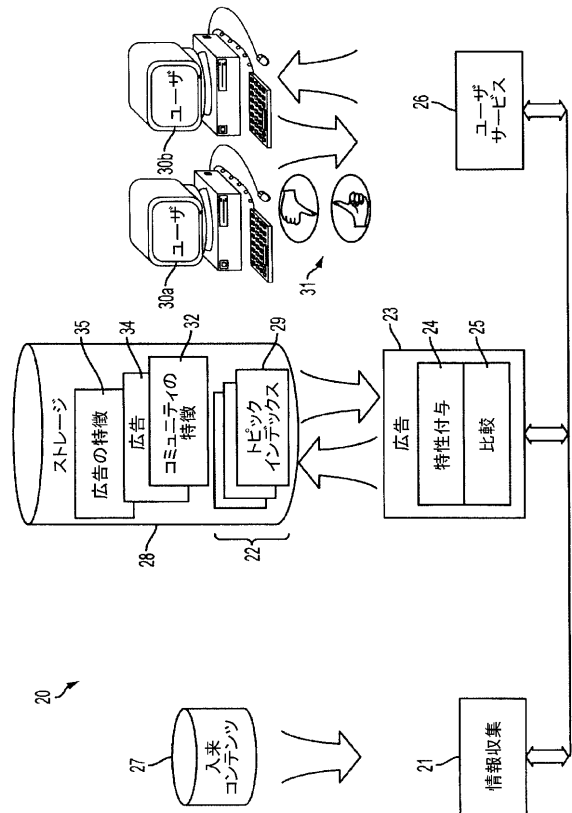


- 5 2 ルータ
- 5 3 単語
- 5 4 単語
- 5 5 n グラム
- 5 6 n グラム
- 6 1 広告
- 6 2 ランディングページ
- 6 3 広告
- 6 4 ランディングページ
- 7 1 表
- 7 2 表

【 図 1 】



【 図 2 】



【 図 3 】

ルータ

From 1980s, the free encyclopedia is usually tailored to the tasks of routing and forwarding. Generally containing a specialized operating system (e.g. Cisco Networks IOS, RAM, NVRAM, flash memory, and one or more processors) high-speed routers are needed to connect and direct a great deal of parallel processing. Circuit-based systems (located right) how multiple ASIC on every module and allow for a wide variety of LAN, WAN, Metro and WAN port. However, with the proper software (such as IOS or Quagga), even commodity PC can act as routers.

Routers connect with two or more logical segments, which do not with router, but switch is really a marketing term without a physical interface types.

Routers operate in two different planes: (2):

- Control Plane, in which the router learns the outgoing interface that is most appropriate for forwarding specific packets.
- Forwarding Plane, which is responsible for the actual process of sending a packet received on a logical interface to an outbound logical interface.

Contents

- 1. Control Plane
- 2. Forwarding Plane (a.k.a. Data Plane)
- 3. Types of Routers
- 4. Small and Home Office (SOHO) connectivity
- 5. Enterprise Routers

ルータ本工

ROUTER: HARD PRODUCT

【 図 4 】

50	特徴的な単語	最高位のNGグラム	ホームルータ	特徴的な単語	最高位のNGグラム	ネットワークルータ
100	ツール	100	パワーツール	100	ジュニバ ネットワーク	ジュニバ ネットワーク
ルータ	34	カクタアウト ツール	34	ネットワーキング	36	ファイアウォール IPセック
アクセサリ	24	ポータ アクセサリ	24	アクセラレイション	33	セキュリティ防外 コパシジョン
マキタ	19	ミルウォーキー ルータ	19	ファイアウォール	32	プロダクト コンパシジョン
コードレス	17	クラフトマン ルータ	17	セキュリティ	30	ネットワーク防外 コパシジョン
ケーブル	16	ポータ カクタアウト	16	ジュノ	29	サービス プロダクト
ポータ	16	ポータ ルータ	16	サーキット	28	ネットスクリーン セキュリティ
ミルウォーキー	15	ポータ ルータ	15	サーキット	27	ルータ シスコ
クラフトマン	15	マキタ ルータ	15	デルタ	26	サービス ソフトウェア
ドリル	14		14	インフラネット	25	
レザーマン	14		14	インターフェース	24	
リコンディショニング	13		13	IPセック	23	
ボッシュ	12		12	インフラネット	22	
マルチ	11		11	カクタアウト	21	
エレクトリック	11		11	ネイラー	20	
シヨップ	10		10	ホイスト	19	
デルタ	10		10	スキル	18	
カクタアウト	9		9	チャット	17	
ネイラー	9		9	ネットワーキング	16	
ホイスト	9		9	モジュール	15	
ヒタチ	9		9		14	
スキル	9		9		13	
チャット	8		8		12	
ネットワーキング	8		8		11	
モジュール	8		8		10	
					9	
					8	
					7	
					6	
					5	
					4	
					3	
					2	
					1	

【 図 5 】

60

61

62

63

64

ホエルターセール

Discount Prices on all Major Brands of Routers, Fast Shipping!  
www.RouterHeadquarters.com

安いシスコルータ

1-Year Warranty, Fast Delivery of top Cisco Router Series Models.  
www.networkliquidators.com

【 図 6 】

70	100	シスコ	72	100
アクセサリ	22.3	ルータ	54	54
デウォルト	18.1	ネットワーク	15	15
マキタ	18	リアーピシジョン	11.7	11.7
ルータ	17.6	セル	8.6	8.6
コードレス	17.5	スウイッチ	6.3	6.3
ケーブル	16.7	アウンドリ	4.8	4.8
ミルウォーキー	15.8	リキディタ	4.8	4.8
クラフトマン	15.5	ノーテル	3.6	3.6
ドリル	14.6	コンパック	3.5	3.5
レザーマン	13.4	GBIC	3.1	3.1
リコンディショニング	12.3	イクストリーム	2.6	2.6
ボッシュ	12.0	シャーン	2.3	2.3
マルチ	11.7	フォラム	2.3	2.3
エレクトリック	11.5	テレコム	2.3	2.3
シヨップ	10.8	クウォート	2.3	2.3
デルタ	10.5	フォース102.3	2.3	2.3
カクタアウト	9.9	M160	2.3	2.3
ネイラー	9.9	オンライン	2.3	2.3
ホイスト	9.9	リアーピシジョン	2.3	2.3
ヒタチ	9.6	9862	2.3	2.3
スキル	9.3	3620	2.1	2.1
チャット	8.2	ネットワーキング	1.5	1.5
モジュール	8.2	モジュール	1.5	1.5

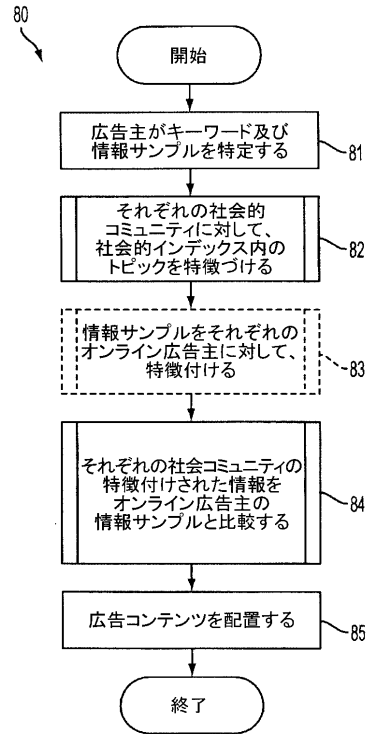
【 図 7 】

AD	木工	ネットワーキング
ネットワークリキデータ	0	3
ルータヘッドクウォータ	23	1

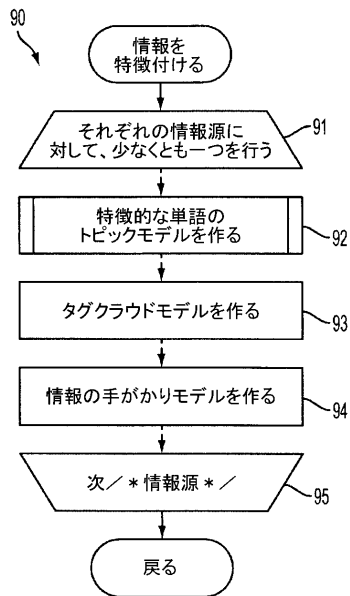
【 図 8 】

AD	木工	ネットワーキング
ネットワークリキデータ	2	8
ルータヘッドクウォータ	25	2

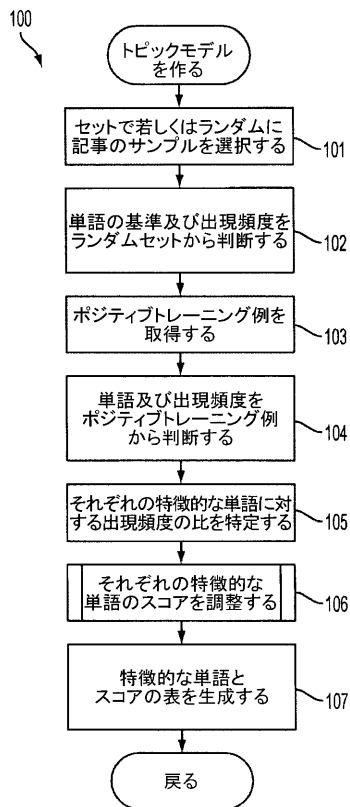
【 図 9 】



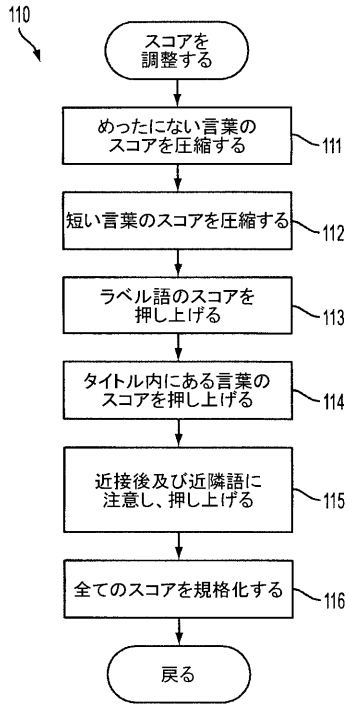
【 図 10 】



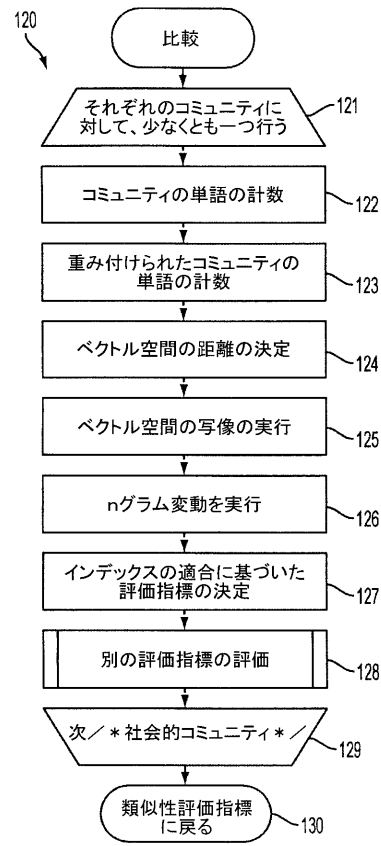
【 図 11 】



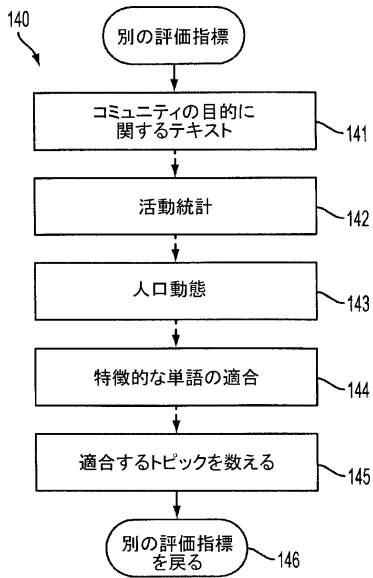
【 図 1 2 】



【 図 1 3 】



【 図 1 4 】



## フロントページの続き

(51)Int.Cl. F I テーマコード(参考)  
G 0 9 F 19/00 Z

(74)代理人 100109335

弁理士 上杉 浩

(74)代理人 100164530

弁理士 岸 慶憲

(72)発明者 マーク ジェイ ステフィック

アメリカ合衆国 カリフォルニア州 9 4 0 2 8 ポルトラ ヴァレー ポルトラ グリーン サ  
ークル 1 0

(72)発明者 ローレンス リー

アメリカ合衆国 カリフォルニア州 9 4 0 2 5 メンロ パーク ハーヴァード アヴェニュー  
7 1 2

(72)発明者 ダニエル エイチ グリーン

アメリカ合衆国 カリフォルニア州 9 4 0 8 7 サニーヴェイル マネ ドライヴ 1 0 5 5  
# 6

(72)発明者 エド エイチ チー

アメリカ合衆国 カリフォルニア州 9 4 3 0 6 パロ アルト グリーンメドレー ウェイ 1 4  
3

F ターム(参考) 5B075 ND20 NK39 QM05