

(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2010-55617

(P2010-55617A)

(43) 公開日 平成22年3月11日(2010.3.11)

(51) Int.Cl.
G06Q 30/00 (2006.01)F I
G06F 17/60 326

テーマコード (参考)

審査請求 未請求 請求項の数 4 O L (全 17 頁)

(21) 出願番号 特願2009-191893 (P2009-191893)
 (22) 出願日 平成21年8月21日 (2009.8.21)
 (31) 優先権主張番号 61/092,736
 (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 12/436,060
 (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 94
 304 パロ アルト カイオーテ ヒル
 ロード 3333
 (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

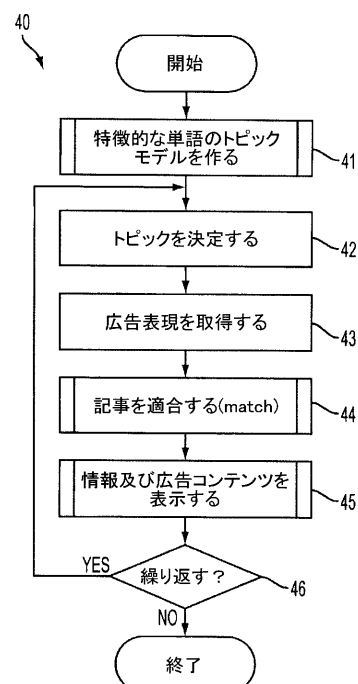
(54) 【発明の名称】 社会的インデックス付けを用いて、トピックに基づいて広告ターゲットを拡大させるシステム及び方法

(57) 【要約】

【課題】 オンライン広告のターゲットに関する問題を解決する。

【解決手段】 デジタル情報の複数の記事、及び、それぞれが1又は2以上の前記記事に関連する複数のトピックを含む1又は2以上のインデックスと、プロセッサ、及び、前記プロセッサのための実行コードを記録したメモリを含むコンピュータであって、1又は2以上の前記記事を含むウェブページを特定し、前記1又は2以上の記事に関連する少なくとも一つの前記社会的インデックス内の前記1又は2以上のトピックを決定する社会的インデックス付けモジュールと、オンライン上の広告主の広告コンテンツに関連付けられた複数の広告表現を受信し、各広告表現を1又は2以上の記事に関連する1又は2以上のトピックに連続的にマッチング（適合）させ、続いて1又は2以上のトピックを記述する単語にマッチング（適合）させる広告処理モジュールと、を含むコンピュータと、を含むことを特徴とするシステム。

【選択図】 図3



【特許請求の範囲】

【請求項 1】

社会的インデックス付けにおいて広告ターゲット拡大をトピックに基づいて行う、コンピュータにより実行されるシステムであって、

デジタル情報の複数の記事、及び、それぞれが 1 又は 2 以上の前記記事に関連する複数のトピックを含む 1 又は 2 以上のインデックスと、

プロセッサ、及び、前記プロセッサのための実行コードを記録したメモリを含むコンピュータであって、

1 又は 2 以上の前記記事を含むウェブページを特定し、前記 1 又は 2 以上の記事に関連する少なくとも一つの前記社会的インデックス内の前記 1 又は 2 以上のトピックを決定する社会的インデックス付けモジュールと、

10

オンライン上の広告主の広告コンテンツに関連付けられた複数の広告表現を受信し、各広告表現を 1 又は 2 以上の記事に関連する 1 又は 2 以上のトピックに連続的にマッチング（適合）させ、続いて 1 又は 2 以上のトピックを記述する単語にマッチング（適合）させる広告処理モジュールと、を含むコンピュータと、

を含むことを特徴とするシステム。

【請求項 2】

前記広告表現は、少なくとも一つの最も密接に適合した広告表現、及び、クリック率によって収益を最大化する広告表現として選択されることを特徴とする請求項 1 に記載のシステム。

20

【請求項 3】

前記社会的インデックス内に備えられる前記トピックは階層的に編成され、

前記広告表現に対するクリック率は、前記 1 又は 2 以上のトピックの範囲を前記階層の連続的なレベルに沿って拡大することによって毛帝されることを特徴とする請求項 2 に記載のシステム。

【請求項 4】

前記クリック率は、複数の前記社会的インデックスにわたって判断され、

前記広告表現に対する前記クリック率は、それぞれの社会的インデックスの前記階層の前記連続的なレベルに加えて、前記社会的インデックスを集めることによって、前記 1 又は 2 以上のトピックの前記範囲を拡大することに応じて、決定されることを特徴とする請求項 3 に記載のシステム。

30

【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、特に社会的インデックス付けを用いて、トピックに基づいて広告ターゲットを拡大させるシステム及び方法にかんする。

【背景技術】

【0002】

広告オークションは、広告主間のほぼリアルタイムの競争にオンライン広告を開放することによって、価格決定の障害を取り除いた。例えば、2001 年 9 月 4 日に発行された米国特許第 6,285,987 号は、計算エージェントを用いて、広告からウェブサイトにキーワードを適合(matching)させることによって、スペースに対して入札することを開示している。この公報は、参照により、本明細書に組み込まれているものとする。同様に、Yahoo! のサービスが提供する、オーバチュア(Overture)は、初めてオンライン検索に広告を統合した。また、オーバチュアは、検索クエリに広告を適合させる入札メカニズムを用いている。それ以来、別のアプローチも発展してきた。例えば、Google Inc. (Mountain View, CA) によって提案されたアドセンス(Ad Sense)は、オークションでの決定に応じて、広告をウェブページの異なる部分に配置する。複数の広告主は、オンライン広告に適合させるのに用いられるキーワードに対して入札し、或いは、広告は、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付

40

50

けられる。

【発明の概要】

【発明が解決しようとする課題】

【0003】

それにも関わらず、効果的にオンライン広告のターゲットを定めることは、困難なままである。例えば、広告をウェブページに載せる目的のキーワードを誤って否定することがあるため、広告をするためのよい機会を逃すことがある。ウェブページ上の情報が、適切な広告を提案するためのキーワードのターゲティング論理キーワードを引き金にすることに失敗するとき、この問題が発生する。言い換えると、広告が、ふさわしい配置であっても、広告を適合させるときに用いられるキーワードを、概念的に関連があるウェブページに合致させることに失敗することがある。したがって、効果的なターゲティングは、ウェブページのコンテンツ以上に考慮しなければならないが、従来の方法では、依然としてうまくいっていない。

10

【0004】

例えば、2001年7月31日に公表された米国特許第6,269,361号は、検索エンジンによって生成された検索結果のリストに影響を与えるシステム及び方法について開示している。この公報は、参照により、本明細書に組み込まれているものとする。ウェブサイトのプロモータは、検索結果のリスト表示のために、検索のリスト化を定義し、プロモータのウェブサイトに関連する検索語を選択し、インターネットの検索エンジン上で検索のリスト化のための位置に影響を与えることができる。その後、ユーザは、検索語を入力するとき、検索エンジンは、ビッド量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る場所に、プロモータのリスト化と共に、検索結果のリストを生成する。しかし、検索語は、検索結果のリストにおいて、提供される特定のウェブコンテンツに強く結びついていることが、予期される。

20

【0005】

2004年3月25日に公開された米国特許公開第2004/0059708号は、関連する広告を提供する方法及び装置を開示している。この公報は、参照により、本明細書に組み込まれているものとする。広告のためのターゲット情報は、ターゲットの記事のコンテンツを分析することによって特定され、トピックのリストを特定する。ターゲット情報をトピックのリストと比較し、広告が、目的の記事に関連することを判断する。しかしながら、入札はトピックに受け入れられるのであるが、トピックは通常、他の誰かに定義されるものであり、エラーが発生することもあり、また、しばしば不正確である。

30

【0006】

2007年11月19日に公表された米国特許公開第2007/0260508号は、階層化した構造によって情報をリスト化する、広告を編成することを開示している。この公報は、参照により、本明細書に組み込まれているものとする。同公報において、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。入札額が、ノードのレベルに応じて送出され、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。階層的な編成及びモードのラベルを、予告なく変更していた可能性もあるが、広告主は入札時にノードと階層のアプローチに切り替えなければならないが、従って、入札を未決定の状態にする。

40

【0007】

また、ユーザの場所、訪問したウェブページの内容、及び検索履歴といったユーザのプロフィール情報が、ターゲティングを向上させるために用いられてきた。ユーザの情報は、ターゲティングプロフィールとして永続的に格納することができる。他の選択肢として、ユーザの興味を調べるために、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も可能である。しかし、昨今、連邦規定に対する申し立てを行っているプライバシーの擁護者が、強調するように、プロフィール情報の利用は、権利を侵害するものだととして反対されている。

【0008】

50

例えば、2001年9月4日に公表された米国特許第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。この公報は、参照により、本明細書に組み込まれているものとする。中央サーバは、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、その他の情報といった、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。

【課題を解決するための手段】

【0009】

キーワードが記事に対して適合し、例を与えることによって、養成される社会的インデックスを用いて、ターゲティングを決定する。一度決定されると、社会的インデックスは、記事をインデックスのトピックのカテゴリに関連付ける。記事をトピックに応じて分類すると、複数の記事から抽出した複数の単語を組み合わせて、キーワードマッチング及び入札のための、より大きな「ターゲット」を作ることができる。したがって、記事のトピックに応じて、いずれかの記事がキーワードに適合すれば、記事がキーワードに適合し得る。

【0010】

一実施形態によって、社会的にインデックス付けを用いて、トピックに応じて広告ターゲットを拡張するための、コンピュータによって実施されるシステム及び方法が提供される。デジタル情報の記事、及び、一又はそれより多い社会的インデックスが維持される。それぞれの社会的なインデックスは、それぞれ一又はそれより多い記事に関連するトピックを含む。一又はそれより多い記事を含むウェブページが特定される。一又はそれより多い記事に関連する少なくとも一つの社会的なインデックス内にある、一又はそれより多いトピックが決定される。オンライン広告主に関するコンテンツを広告することにそれぞれ関連する複数の広告表現を受け取る。それぞれの広告表現は、一又はそれより多いトピックを表現する言葉に適合することによって導かれる一又はそれより多い記事に関する一又はそれより多いトピックに連続的に適合する。

【0011】

別の実施形態は、社会的インデックス付けを用いて、トピックに応じて広告ターゲットを拡張するための、システム及び方法を提供する。一連のトピックは、オンラインコミュニティインデックス内で、維持される。また、それぞれのインデックストピックに対する一連の最も特徴的な言葉が、維持される。広告サービスは、オンライン情報のリソースとしての場所に対する広告とキーワードを提供する。広告サーバは、トピックエンジンを提供する。そのトピックエンジンは、広告のキーワードをコミュニティのインデックストピックに関連する特徴的な単語に適合させることによって、オンラインコミュニティのユーザに広告を向ける。

【0012】

実施形態は、従来の方法を超えた多くの利点を提供する。第一には、トピックの下にある複数の記事を、ターゲティング広告のために、考慮することができるので、原理的に誤って否定することが減少する。第二には、プライバシーの侵害が起きない。なぜなら、サイトを移動するユーザを追跡せず、ユーザの個人的な閲覧趣向を調査する必要がないからである。代わりに、このシステムは、個人的な振る舞いに関する情報よりもむしろ、コミュニティに関するデータ及びトピックに関するデータを、ターゲティングのために用いる。

【0013】

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、意向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

10

20

30

40

50

な説明は、例示目的であり、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】

【0014】

【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】一実施形態に適合した社会的なインデックス付けを用いて、トピックに応じて、広告ターゲットを拡大する方法を示す。

【図4】図3のルーティーンを用いる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5】図4のルーティーンを用いる特徴的な単語のスコアを調整するルーティーンを示すフロー図である。

【図6】図3の方法を用いて、記事に広告表現を適合するルーティーンを示すフロー図である。

【図7】図3の方法を用いて、記事にオンライン広告を表示するルーティーンを示すフロー図である。

【図8】記事及び社会的インデックスと共に提供されるオンライン広告を用いるユーザインターフェースのスクリーンショットを、例として、示す図である。

【図9】テーブル型に編成した記事を用いるユーザインターフェースのスクリーンショットを、例として、示す図である。

【発明を実施するための形態】

【0015】

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し、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し、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11及びトピック検索システム12が、連携して機能し、意味付け及び検索をサポートする。

【0016】

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コーパス(corpus)である。情報の量に制約はなく、そのことはコーパス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16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。

【0017】

ネットワーク16は、多様で分散された情報源とコンシューマとの相互接続を可能にし、前者はコーパスにデータ投入し(populate)、後者はこれにアクセスする。著者、編者、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16に掲載する。これらは、ウェブサーバ14a、ニュースの集合サーバ14b、投票可能なニュースサーバ14c、その他の情報ソースを通して、分配されるデータのコーパスとして維持される。これらのソースは、それぞれウェブコンテンツ15a、ニュースコンテンツ15b、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15cを、ネットワーク16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13a-cを介して行われる。サーバその他の非ユーザデバイス情報コンシューマ機器も同様にコーパス内に維持された情報を調査、検索、使用することはできるが、ここでは明確にするために、ユーザのデバイスだけを考える

10

20

30

40

50

。

【0018】

一般に、ユーザデバイス13a-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、ウェブが利用可能なデバイスであり、サーバ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14a-cを用いて、情報の交換及び検索をサポートする。ユーザデバイス13a-c及びサーバ14a-c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、不揮発性ストレージ等を含んでいる。ただし、これら以外のものを含んでもよい。さらに、サーバの代わりに、若しくは、サーバ14a-cに加えて、別の情報ソースを含んでもよいし、ユーザデバイスの代わりに、若しくは、ユーザデバイス13a-cに加えて、別の情報コンシューマ機器を含んでいてもよい。

10

【0019】

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12は、社会的インデックス付けシステム11に統合される。社会的インデックス付けシステム11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12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、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、ユーザデバイスからの検索クエリは、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12によって、サーチ結果を表示するために生成される。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。その上、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観点を利用して、広告を検索された情報に結び付けるために用いられるターゲティングの範囲を広げることによって、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。

20

【0020】

デジタル情報の検索のための環境10は、ユーザからは単一の情報ポータルとして見えるが、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10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、探査(prospecting)、適応(orienting)のような別の関連した動作を行うために、追加の要素、若しくは、機能モジュールを要求してもよい。

30

【0021】

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、ユーザサービス26の各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コーパスから継続的に取り入れら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。記事の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。

40

50

【0022】

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、コーパス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、同一出願人によって2008年8月12日に出願された一連番号12/190, 550の米国特許公開のように、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ーグリーン（evergreen）インデックスとすることができる。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。エバー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の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、エバー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。

10

【0023】

以下で図3に関連してさらに説明されるように、広告モジュール23は、オンライン広告のターゲティングの幅を広げる。広告モジュール23は、各トピックインデックス29の中のトピックについて、特徴的な単語のトピックモデルとしても知られる粗いトピックモデル32を作る、予め計算するサブモジュール24を含む。この広告モジュールは、検索エンジン（非表示）に結び付いたプロセッシングサブモジュール25も含む。プロセッシングサブモジュール25は、記事自身の内容、若しくは、インデックスに含まれる記事のトピックに関連する情報に基づいて、広告コンテンツ34を特定した記事に適合させる。次いで、記事、及び、それに最も密接に適合する広告が、ユーザインターフェースを介してユーザに視覚的に表示される。別の実施形態においては、2009年5月5日に出願した一連番号12/436, 067の同一出願人による米国特許出願に記載されたように、オンライン広告は、特定の社会的コミュニティにターゲットを絞り、そのコミュニティのトピックは、特定の広告主の興味に最も合ったものを反映している。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。それぞれのトピックインデックス29及び一連のトピックのインデックス22に関する粗いトピックモデル32は、集約されたストレージ28に、保持される。

20

30

【0024】

最後に、ユーザサービスモジュール26は、フロントエンドをユーザ30a-bに提供し、一連のトピックのインデックス22及び入来コンテンツ27にアクセスし、一連のトピックのインデックス22、若しくは、一つのトピックのインデックス29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29は、中心的な課題領域に対して継続的な関心を持つ、「増大した(augmented)」コミュニティとも呼ばれるユーザのコミュニティに結び付けられる。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30ことによって引用される情報を「検査(vet)」する。

40

【0025】

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して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が、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）プライスモデルの基づいた計算の基礎となる。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が、通常は、CPC（クリック当たりのコスト）プライスモデルに基づいた収入計算のための基礎となる。三つ目は、サイト上で登録したり、商品を購入したりするといった、ユーザが、広告主のサイトに対して行動を取るときに発生するコンバ

50

ージョン(conversion)イベントである。一般に、広告主は、コンバージョンイベントに関する統計を管理しており、期待されるコンバージョンのレート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。

【0026】

適切なマッチング(matching)コンテンツに対するターゲティング、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。降下のないターゲティングによっては、収入を得る機会が、失われるかもしれない。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、ターゲティングを改良し、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テーマの領域に記事を表示し、ユーザが見つけたトピックに関する情報にすばやくアクセスするのに役立つ。その上、社会的インデックスは、多くの利用者が興味を持つ記事を編成するコミュニティによって利用される。従って、社会的インデックスは、オンライン広告に用いて、ユーザをコミュニティに分類する情報サービスに広告のターゲットを定めることができる。これにより、潜在的な広告収入を増加させながら、ユーザを特定するといった行為を避けることができる。

【0027】

広告は、トピック検索の結果と共に、あるいは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ス上に存在する、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く。そして、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ターゲティングをした広告になる。このような情報への道は、直接的な記事検索を介することになり、それは、情報又はトピックの軌跡に従う結果としてのトピック検索を実行することによって、あるいは、情報を求めてアクセスする別の方法によって行われる。

【0028】

従って、ターゲティング及び広告の配置をより効果的にするために、オンライン広告を、共有されたトピックの社会的なインデックスから得られるコミュニティ情報を用いて、オンラインコミュニティに向けて発することができる。それぞれの社会的インデックスは、同じインデックストピックに分類されたトピック的に関連した記事全体から、コンテンツの集計分析を行う。従って、任意のウェブページ若しくはユーザのプロフィールに存在する言葉のみに基づいて広告をターゲティングするのではなく、トピックごとに分類したインデックスによって一つの記事にターゲットされた広告の幅を広げることができる(そうすることが適切な場合に)。図3は、一実施形態に従った、社会的なインデックス付けにおける広告ターゲットのトピックに基づいて拡大させる方法40を示す。この方法40は、一連の処理若しくは方法のステップとして、例えば、サーバのような汎用目的でプログラムされたコンピュータによって、実行される。

【0029】

このアプローチは、社会的コミュニティ情報を用いることで、ターゲティングの際誤った否定語を減らすことによって、広告収入を改善させる。方法40は、二つの段階に分けて実行される。第一段階では、図4及び5に示されるように、予め計算している段階の間に、粗いトピックモデル、すなわち、特徴的な単語のトピックモデルが、それぞれの社会的インデックスにおいてリスト化されたトピックに応じて作られる(ステップ41)。それぞれの粗いトピックモデルは、それぞれのトピックに対して重要な言葉、若しくは、それぞれのトピックを反映している言葉を特定する。粗いトピックモデルは、特定されたそれぞれの特徴的な単語に対して、相対的な重要性を示す計量若しくはスコアを含んでいる。

【0030】

第二段階では、オンライン広告の要求が処理される（ステップ４２－４６）。広告を要求する形式がどのようなものかは、重要ではなく、競争によるものでも、直接購入であってもよい。例えば、この要求は、広告場所を決めるためのオークションスタイルの形式による競争入札でもよいし、広告に対する現在の要求を集めたプールから選択するものでもよいし、あるいは広告の別の形式でもよい。情報の要求（ステップ４２）によって関連付けられるトピックは、それが特定のコンテンツ又は検索結果に関係するトピックであるか、トピックツリーのセクションにおいて見られるようなより一般的なトピックであるかに関係なく、広告コンテンツの位置をガイドするためのターゲットとなる。例えば、図８及び図９を参照して以下にそれぞれ示すように、記事及び社会的インデックス若しくは表形式で編成された記事を含むウェブページに関連したトピックに、ターゲティングを向けることができる。それぞれのオンライン広告は、ウェブページ上に配置され表示されるコンテンツと、広告表現とを含む。その広告表現は、コンテンツを適切な記事にターゲットを合わせ、マッチングさせるために用いられている。広告表現は、一般には一連のキーワードであり、語幹、同義語、及び、一般化された別の形式を含むことができる。図３に関連して以下に示されるように、オンライン広告それぞれに対する広告表現は、最初に検索され（ステップ４３）、次いで、要求された情報に関するトピックにマッチングされる（ステップ４４）。図５に関連して、以下に示されるように、最も密接に適合した広告表現が特定され、情報及び広告コンテンツが表示される（ステップ４５）。オンライン広告要求の処理段階（ステップ４２－４６）は、それぞれの情報要求に対して繰り返し実行される（ステップ４６）。

10

20

【００３１】

予備計算段階を、オンライン広告要求処理とは独立して先行して実行することができる。粗いトピックモデルは、予備計算段階の間に生成され、これらのモデルは適切なときに、オンライン広告のターゲティングの幅を広げるために用いられる。図４は、図３のルーティーン４０と共に用いる粗いトピックモデルを作るルーティーン５０を示すフロー図である。粗いトピックモデルには、特徴的な複数の単語と、それぞれの特徴的な単語の相対的重要性を反映するスコアが含まれる。

【００３２】

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ものであり、これらは、典型的には適切なコーパス内の記事から選択された言葉である。そのコーパスは、ウェブページ、電子書籍、若しくは、その他の印刷資料として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、コーパスから選択する（ステップ５１）。特徴的な単語の基準及び発生頻度が、選択した記事から抽出される（ステップ５２）。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のコーパス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、個別のインデックス２９の全てに対するコーポラ(corpora、corpusの複数形)である、システム全体のコーパス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量される。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だけ、変化した場合には、存在の頻度が再度決定される。

30

40

【００３３】

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例が得られる（ステップ５３）。ポジティブトレーニング例は、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５４）。計量若しくはスコアが、例えば、語頻度－逆文献頻度（ＴＦ－ＩＤＦ）の重みを用いて、それ

50

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内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55）。図4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56）、トピックに対して特徴的な単語の重要度を向上させるか、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57）。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。

【0034】

異なる「アクター（actor）」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。人間であるインデックス管理者は、ポジティブトレーニング例を選択する。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、コーパス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コーパス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。

【0035】

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5は、図4のルーティーン50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60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、コーパス、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61）。同様に、文字数が最小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62）。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63とステップ64それぞれ）。典型的には、全てのラベルの言葉は、特徴的な単語として含められる。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(proximal)」語のスコアは、押し上げられる（ステップ65）。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66）。最高のスコアを持つ特徴的な単語はまた、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、コーパス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。

【0036】

オンライン広告は、広告表現を用いて、ユーザによって特定されたコンテンツとマッチング。図6は、図3の方法40を用いて、記事に広告表現を適合するルーティーン70を

示すフロー図である。それぞれの広告表現は、繰り返し処理される（ブロック71ー76）。マッチングは、二つのフェーズで行われる。最初にそれぞれの広告表現、一般には広告内容を表現するキーワードを、ユーザによって要求された情報に直接的にマッチングしようとする（ステップ72）。その情報は、特定の記事、選択したウェブページ内にある一連の記事、及び、一連の検索結果などであってよい。広告は、適合した場合には（ステップ73）、配置される対象とみなされ（ステップ77）、要求処理が完結する。

【0037】

しかしながら、広告表現が、要求された情報に表現上でマッチング（適合）しない場合には、社会的インデックスの助けを借り、ユーザの所属する社会的コミュニティを用いてターゲティングを幅広くする。それぞれの広告表現を、社会的インデックス内にある要求された情報のトピックに関連した記述的な言葉に間接的に適合させようとする（ステップ74）。要求された情報が、ウェブページ上に表示され、ウェブページを介して、ユーザが所属する社会的コミュニティに関する社会的インデックス内にあるトピックに、広告要求を適合させる。ユーザの意向は、トピックを幅広いものにする際のガイドとして、ユーザに表示される情報は、意向に対する信号を与える。例えば、特定の記事若しくは一連の記事を検索するというユーザの要求、若しくは、検索クエリのようなより一般化されたユーザの要求から意向を判断することができる。それぞれの記事、若しくは、検索クエリの結果が、社会的インデックスに応じて、トピックにマッチングする。同様に、意向は、ユーザに関する社会的インデックスに対するトピックツリーの拡大した部分を反映する。その部分は、ユーザの示唆した特定の一連のトピックに関する興味を示す。ユーザの意向を表す別の信号も可能である。

【0038】

広告要求は、推論したトピックを記述する単語にマッチングされる。例えば、記述的な言葉は、トピックに応じてインデックスが付けられたそれぞれの記事内にある、全ての単語を含むことができる。その単語は、特定のウェブページ上の記事に対応しているか、若しくは、トピックに応じて決定された特徴的な単語であってもよい。社会的インデックスに対してトピックツリーの選択した部分によって示唆されるように、ユーザの意向をより幅広い範囲でとらえる場合には、特徴的な単語は、選択したトピックを基に判断される。別のトピックに応じた記述的な言葉を別の形式でとらえることも可能である。マッチングされる（適合される）場合には（ステップ75）、広告は、場所を与えられる候補とみなされる（ステップ77）。処理は、残りの広告表現それぞれに対して続けられる（ステップ76）。

【0039】

空間的な制約のため、限られた数のオンライン広告のみを、ユーザによって要求された情報と共に示すことができる。図7は、図3の方法40を用いて、記事にオンライン広告を表示するルーティーン80を示すフロー図である。好ましくは、最も適した広告が、広告モジュール23（図2に示す）若しくは別のサーバによって選択される（ステップ81）。最も適した広告とは、要求した情報自身、若しくは、社会的インデックス内にある要求した情報のトピックに関連した記述的な単語に、最も密接に適合するオンライン広告であるかもしれない。

【0040】

別の実施形態においては、最も適した広告は、適合したキーワードだけではなく、統計的なクリック率を考慮することによって選択される。この目的は、ユーザによってほとんど選択されず、「クリックされも」しない高い入札額の広告を無視することにある。このシステムは、ユーザが広告の上をクリックする頻度の統計情報を、所持し、広告の入札額に対するクリック率の推量をかけることによって、広告による収入を最大化するために、最も適した対象の広告を選択することができる。また、トピックは、広告選択の因子にもなり得る。親トピックは、次に大きなターゲット及びユーザコミュニティとして任意のトピックを包含する。広告に関する統計、若しくは、広告の関連する集合に関する統計は、コミュニティを中心としたそれぞれの範囲において判断される。システムが、ターゲット

の範囲を拡大するごとに、統計の対象範囲も広がる。表示すべき広告を選択するときには、収入を最大化する広告を選択する前に、一連の範囲に対して、競合する複数の広告を考慮する。

【0041】

多くの場合、広告キャンペーンを始めるときには、クリック率の推量を基にしたデータを利用することはできない。キャンペーンを開始するときにクリックが散発的に発生することがあるが、それは広告に最も適する場所として現在考慮したウェブページと必ずしも同じウェブページではない。選択処理を起動してクリックの推量データを提供するために、複数のページからの統計を集めることができる。例えば、最初は、全ての社会的インデックスにわたって統計を集めることができる。統計が蓄積されるにつれて、全てのトピックにわたって集めるのではなく、ウェブページの属する社会的インデックスに対してのみ統計を集めるようにすることができる。統計がさらに蓄積されていけば、トップレベルのトピックについて、統計を集めることができる。継続して行い、それぞれの段階において、統計が集まるにつれて、集める対象領域は狭くなる。最終的には、統計を個別のトピックに対して集めることができる。おそらく、最も狭いときに、統計は最も正確であろう。

10

【0042】

最も適した広告を選択するための別の基準を選んでもよい。さらに別の実施形態においては、上位5つの広告を選択するように、最も密接に適合する広告候補が複数選択される。

【0043】

第三者の広告配置サーバのような広告モジュール23若しくは別のサーバによって、要求された情報の中での配置が判断される。オークションスタイルの形式を用いて、広告主が、競い合う場合には、コストモデルの入札額によって配置を割り当ててもよい。例えばC P Mコストモデルによって、特定の記事に並べて、広告するコンテンツを配置してもよいが、C P Cコストモデルによって、広告するコンテンツをその記事に並べて、若しくは、社会的インデックス内で特定されたトピックの近くに配置してもよい。同様に、紺バージン当たりのコスト(cost-per-conversion)のコストモデルによって、広告するコンテンツを、記事を見るときの視野の中に配置してもよい。コストモデル及び広告するコンテンツの配置を、別の形式にすることも可能である。最終的には、記事及び広告が表示される(ステップ83)。

20

30

【0044】

社会的インデックスの文脈では、広告する内容を表示して、トピックに応じた編成、及び、情報のコミュニティに応じた表示をいっそう有利に活用することができる。図8は、記事91及び社会的インデックス92と共に提供されるオンライン広告を用いるユーザインターフェースのスクリーンショット90を、例として示す図である。記事91は、ユーザによって要求された特定の情報を表現するが、社会的インデックス92は、記事91に関連する、付随する階層的なトピックの編成スキームを供給する。ディスプレイ全体の中に、記事91の付近94、若しくは、社会的インデックス92自身の内側95に配置された広告するコンテンツと共に、従来からあるバナー広告93を表示してもよい。

【0045】

社会的インデックスに応じた情報の別のレイアウト若しくは編集も可能である。例えば、図9は、テーブル型に編成した記事104を用いるユーザインターフェースのスクリーンショット100を、例として示す図である。ユーザインターフェースは、四つのユーザのコミュニティ、「技術ニュース」「開発者」「フリーク」及び「消費者7」それぞれに対して、ユーザが、選択可能なタブ101 a-dを含む。これらのコミュニティそれぞれが、コミュニティによって特定された、異なった視点及び興味に関連した社会的インデックスを持つ。それぞれのタブの内側に、記事104が、社会的インデックスに関するトピックに対応する列内で編集される。例えば、「技術ニュース」のタブは、カリフォルニアのクパチーノにあるアップル社によって製造されたiPhone及びスマートフォンについての記事104の要約を含んでいる。記事の「iPhone」クラスは、階層的なスキ

40

50

ーム、ここでは、スーパークラス102「携帯電話」及び「Sci/Tech」それぞれの下で編成される。また、「iPhone」クラスは、サブトピックス103a-c「アプリケーション」「ニューモデル」「レビュー」に加えて、それぞれのトピックに関する「wiki」の記事105を表示する。iPhone wikiの記事は、iPhoneの要約であり、コミュニティのメンバーによって書かれたものである。それぞれのコミュニティに応じて編成されたタブによって提供される、より充実したコンテンツに基づいて、オンライン広告が、表示される記事の内容と共に、コミュニティ及びトピックに応じて、ターゲットされる。

【0046】

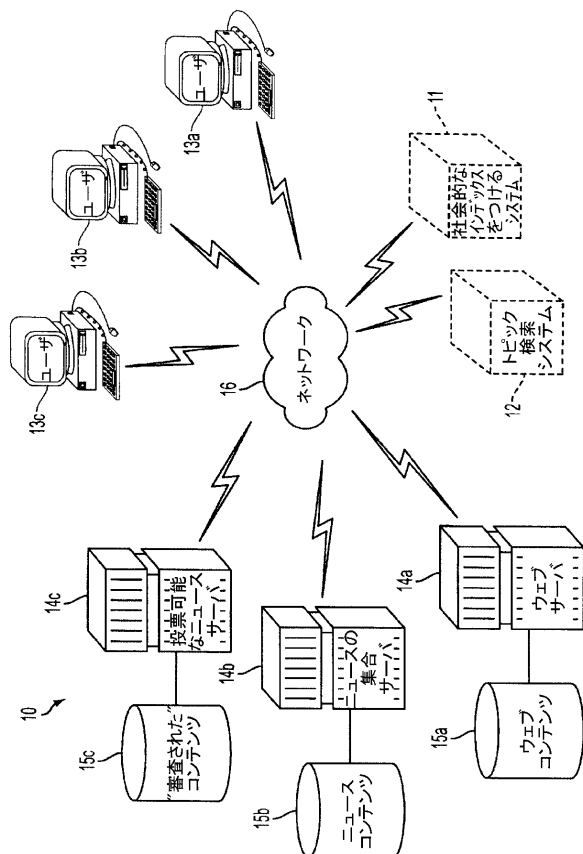
本発明が、特にその実施形態を参照して示されて記載されているが、本分野の当業者は、形式及び詳細において、本発明の思想及び範囲を外れることなく、既に述べたあるいはその他の変更を行うことができることを理解するであろう。 10

【符号の説明】

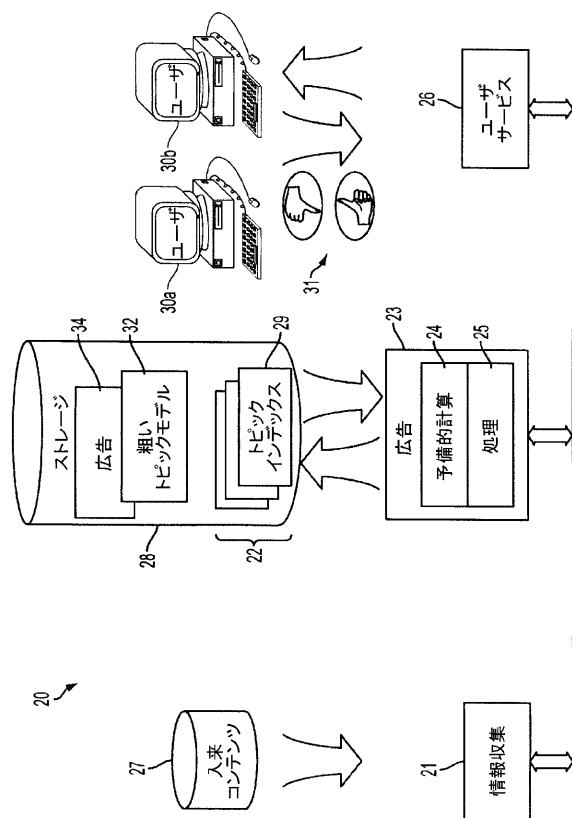
【0047】

10	環境	
11	システム	
12	トピック検索システム	
13a-c	ユーザデバイス	
14a	ウェブサーバ	
14b	ニュースの集合サーバ	20
14c	投票可能なニュースサーバ	
15a	ウェブコンテンツ	
15b	ニュースコンテンツ	
15c	“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	
22	一連のトピックのインデックス	
23	広告	
24	予め計算するサブモジュール	30
25	プロセッシングサブモジュール	
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	粗いトピックモデル	
33	トピックインデックス	
34	広告コンテンツ	
90	スクリーンショット	40
91	記事	
92	社会的インデックス	
93	バナー広告	
94	記事91の付近	
95	社会的インデックス92自身の内側	
100	スクリーンショット	
101a-d	タブ	
102	スーパークラス	
103a-c	サブトピックス	
104	記事	50

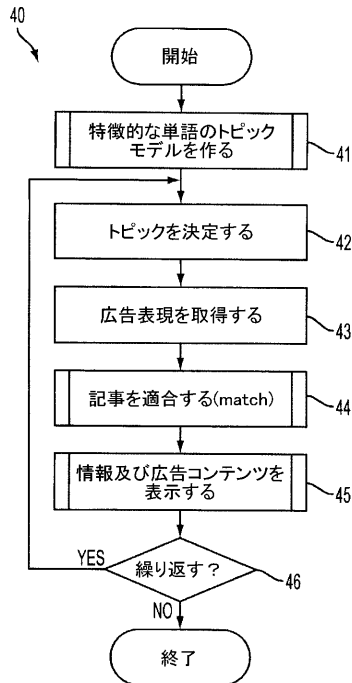
【図 1】



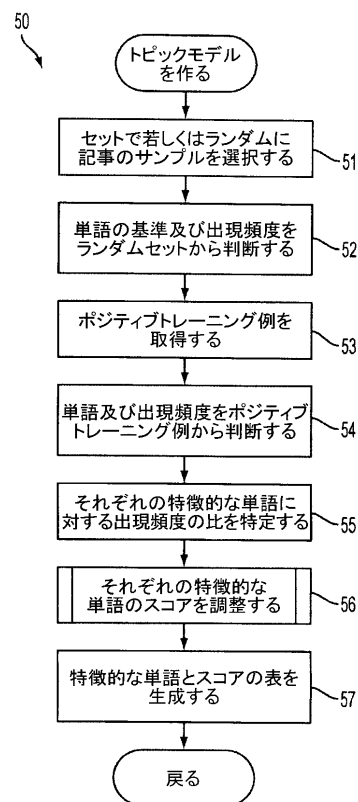
【図 2】



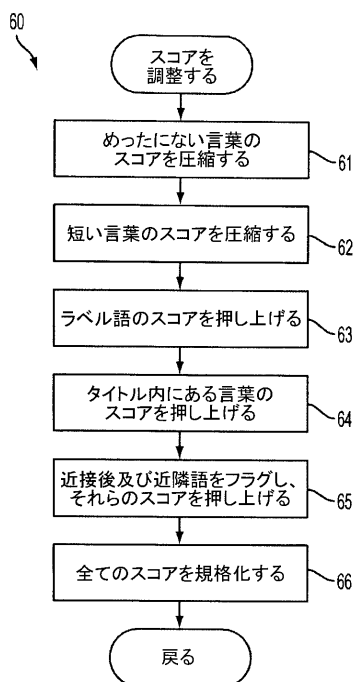
【図 3】



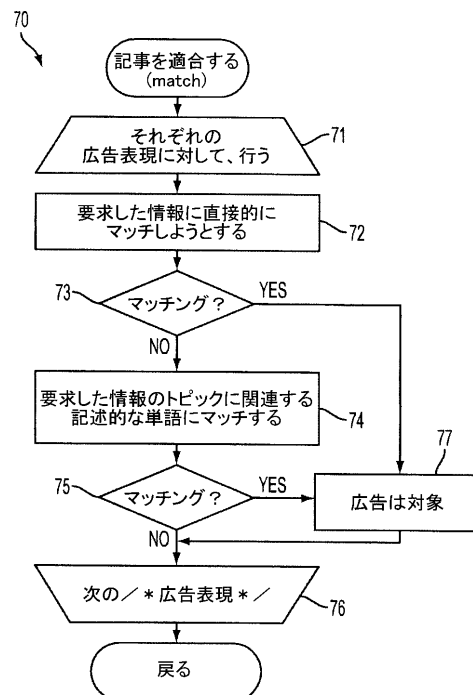
【図 4】



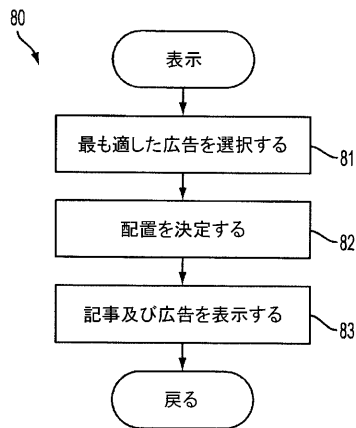
【図 5】



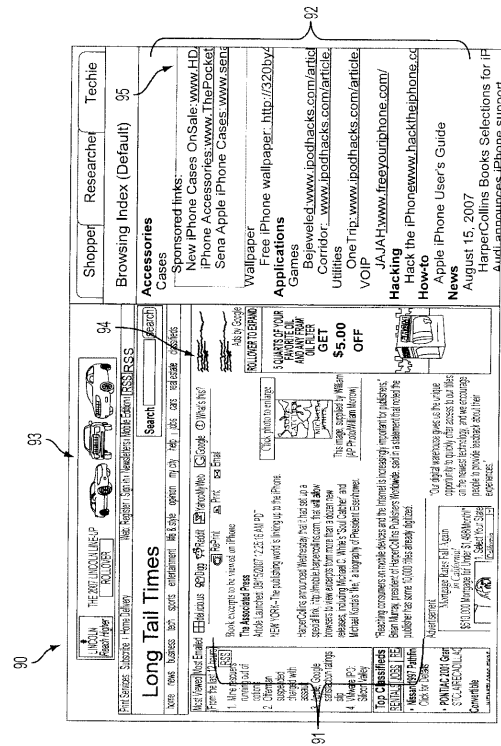
【図 6】



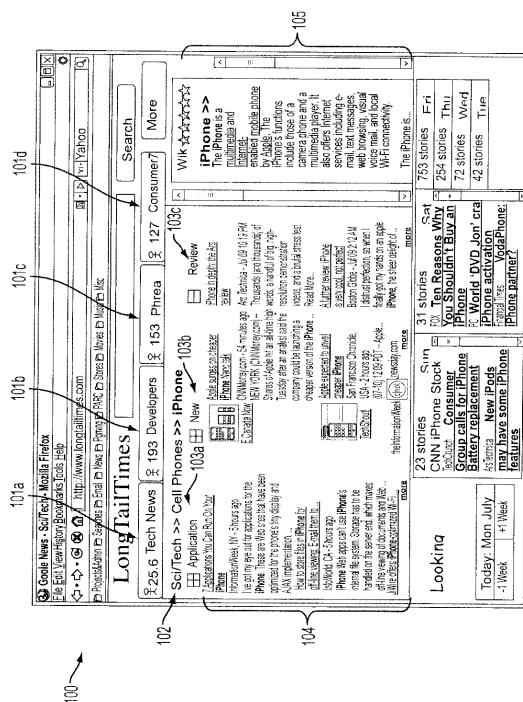
【图 7】



【图 8】



【圖 9】



フロントページの続き

(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 94028 ポルトラ ヴァレー ポルトラ グリーン サー
クル 10

(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 94025 メンロ パーク ハーヴァード アヴェニュー
712

(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 94087 サニーヴェイル マネ ドライヴ 1055
#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 94306 パロ アルト グリーンメドレー ウェイ 14
3