

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5513624号
(P5513624)

(45) 発行日 平成26年6月4日 (2014.6.4)

(24) 登録日 平成26年4月4日 (2014.4.4)

(51) Int.Cl.

F I

G 0 6 F 17/30 (2006.01)

G 0 6 F 17/30 3 4 O Z

G 0 6 F 17/30 2 1 O D

請求項の数 18 (全 26 頁)

(21) 出願番号	特願2012-530856 (P2012-530856)	(73) 特許権者	510330264
(86) (22) 出願日	平成22年9月8日 (2010.9.8)		アリババ・グループ・ホールディング・リミテッド
(65) 公表番号	特表2013-506189 (P2013-506189A)		ALIBABA GROUP HOLDING LIMITED
(43) 公表日	平成25年2月21日 (2013.2.21)		英国領、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・オー・ボックス 847
(86) 国際出願番号	PCT/US2010/002456		
(87) 国際公開番号	W02011/037603	(74) 代理人	110000028
(87) 国際公開日	平成23年3月31日 (2011.3.31)		特許業務法人明成国際特許事務所
審査請求日	平成25年7月25日 (2013.7.25)	(72) 発明者	リン・フェン
(31) 優先権主張番号	200910175066.3		中華人民共和国 ハンチョウ、ワーナー・ロード、ウエスト・レイク・インターナショナル・プラザ、10階、ナンバー391
(32) 優先日	平成21年9月27日 (2009.9.27)		最終頁に続く
(33) 優先権主張国	中国 (CN)		
(31) 優先権主張番号	12/807,505		
(32) 優先日	平成22年9月7日 (2010.9.7)		
(33) 優先権主張国	米国 (US)		
早期審査対象出願			

(54) 【発明の名称】 クエリの一般属性に基づく情報の検索

(57) 【特許請求の範囲】

【請求項 1】

コンピュータによって実行される情報を検索するための方法であって、

クライアントから現行クエリデータを受信し、

前記現行クエリデータの一般属性特徴を抽出し、前記一般属性特徴は、複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテゴリに属する確度を示し、

前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の確度を決定することであって、

前記一般属性特徴を条件付き確率モデルの所与情報として設定し、

前記条件付き確率モデルを用いて、前記現行クエリが前記所与情報の条件下で前記それぞれのカテゴリに属する複数の事後確率を計算し、

前記複数の事後確率を前記複数の確度として設定すること、を備え、

前記複数の確度に少なくとも部分的に基づいてカテゴリを選択し、選択されたカテゴリである選択カテゴリは前記複数のカテゴリの1つであり、他のカテゴリの確度よりも高い確度を有し、

前記選択カテゴリ内において、前記現行クエリデータに対応する検索結果を検索し、

前記検索結果を返すこと、

を備える、方法。

【請求項 2】

請求項 1 に記載の方法であって、前記一般属性特徴は、前記現行クエリデータ内のキーワードに対応する単語レベル特徴を含む、方法。

【請求項 3】

請求項 1 に記載の方法であって、前記一般属性特徴は、前記現行クエリデータ内の見出し語に対応する語根に対応する構文特徴を含む、方法。

【請求項 4】

請求項 3 に記載の方法であって、前記構文特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータ内の前記キーワードの品詞タグ付けと、前記現行クエリデータ内の前記キーワードを含む句の特定とを実行し、

前記品詞タグ付けおよび前記句の前記特定の結果に少なくとも部分的に基づいて、前記クエリデータ内の名詞句を取得し、

前記名詞句の前記見出し語に関連する前記語根を抽出すること、
を備える、方法。

【請求項 5】

請求項 1 に記載の方法であって、前記一般属性特徴は、前記キーワードに対応する意味タグに対応する意味特徴を含む、方法。

【請求項 6】

請求項 5 に記載の方法であって、前記意味特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータの前記キーワードの種類を特定し、
前記特定された種類に従って、前記キーワードの動作動詞タグ、製品タグ、ブランドタグ、型式タグ、または、別のタグを取得すること、
を備える、方法。

【請求項 7】

請求項 1 に記載の方法であって、前記一般属性特徴は、クリックログ内で前記現行クエリデータに対応する被クリックカテゴリに対応するクリック特徴を含む、方法。

【請求項 8】

請求項 7 に記載の方法であって、前記クリック特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータが前記クリックログ内に存在するか否かを判定し、前記クリックログは、前記クエリデータおよび前記クリック特徴の間の対応関係を格納するよう適合されており、

前記現行クエリデータが前記クリックログ内に存在する場合に、前記事前設定されたクリックログから前記現行クエリデータに対応する第 1 の所定の数の被クリックカテゴリを取得し、前記被クリックカテゴリは、前記クエリデータの検索後にユーザによってクリックされた対応する製品が属するカテゴリであり、

前記現行クエリデータが前記クリックログ内に存在しない場合に、前記現行クエリデータに最も近い被クリッククエリデータについて前記クリックログに照会し、前記被クリッククエリデータに対応する前記第 1 の所定の数のカテゴリを前記クリック特徴として取得し、前記被クリッククエリデータは、前記現行クエリデータから最短編集距離にある前記クリックログ内のデータであること、
を備える、方法。

【請求項 9】

請求項 1 に記載の方法であって、前記一般属性特徴は、選択ログ内で前記現行クエリデータに対応する少なくとも 1 つの被選択カテゴリに対応する選択特徴を含む、方法。

【請求項 10】

請求項 9 に記載の方法であって、前記選択特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータが前記選択ログ内に存在するか否かを判定し、前記選択ログは、前記クエリデータおよび前記選択特徴の間の対応関係を格納するよう適合されており、
前記現行クエリデータが前記選択ログ内に存在する場合に、前記選択ログから前記現行クエリデータに対応する第 1 の所定の数の被選択カテゴリを取得し、前記被選択カテゴリ

10

20

30

40

50

は、検索エンジンのマッピングテーブル内で前記現行クエリデータに関連付けられたカテゴリを表し、

前記現行クエリデータが前記選択ログ内に存在しない場合に、前記現行クエリデータに最も近い被選択クエリデータについて前記選択ログに照会し、前記被選択クエリデータに対応する前記第1の所定の数のカテゴリを前記現行クエリデータの選択特徴として取得し、前記被選択クエリデータは、前記現行クエリデータから最短編集距離にある前記選択ログ内のデータであること、

を備える、方法。

【請求項11】

請求項1に記載の方法であって、前記一般属性特徴は、同一セッションにおける前記現行クエリデータの前の1つのクエリに対応する実質的に最適なカテゴリに対応するセッション特徴を含む、方法。

10

【請求項12】

請求項11に記載の方法であって、前記セッション特徴を含む前記一般属性特徴の抽出は、

前記同じセッションにおける前記現行クエリデータの前の先行するクエリデータを取得し、

前記先行するクエリデータに対応する最適なカテゴリ分け結果が属するカテゴリを取得し、

前記最適なカテゴリ分け結果が属する前記カテゴリを前記現行クエリデータのセッション特徴として設定すること、

20

を備える、方法。

【請求項13】

請求項1に記載の方法であって、さらに、

前記複数の確度の内、既定の第1の閾値を満たす選択された確度と、前記選択された確度に対応するカテゴリとを特定し、

前記現行クエリデータについての前記選択カテゴリを選択するために、前記複数の確度の内の前記選択された確度と、前記選択された確度に対応するカテゴリとを前記クライアントに返すこと、

を備える、方法。

30

【請求項14】

請求項1に記載の方法であって、前記選択カテゴリは、前記現行クエリデータについて計算された最大確度を有する、方法。

【請求項15】

情報を検索するためのシステムであって、

1または複数のプロセッサであって、

クライアントから現行クエリデータを受信し、

前記現行クエリデータの一般属性特徴を抽出し、前記一般属性特徴は、複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテゴリに属する確度を示し、

40

前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の確度を決定することであって、

前記一般属性特徴を条件付き確率モデルの所与情報として設定し、

前記条件付き確率モデルを用いて、前記現行クエリが前記所与情報の条件下で前記それぞれのカテゴリに属する複数の事後確率を計算し、

前記複数の事後確率を前記複数の確度として設定すること、を備え、

前記複数の確度に少なくとも部分的に基づいてカテゴリを選択し、選択されたカテゴリである選択カテゴリは前記複数のカテゴリの1つであり、他のカテゴリの確度よりも高い確度を有し、

50

前記選択カテゴリ内において、前記現行クエリデータに対応する検索結果を検索し、
前記検索結果を返すことを実行するよう構成されている１または複数のプロセッサと、
前記１または複数のプロセッサに接続され、前記プロセッサに命令を提供するよう構成
されている１または複数のメモリと、
を備える、システム。

【請求項１６】

情報を検索するためのコンピュータプログラムであって、
クライアントから現行クエリデータを受信するための機能と、
前記現行クエリデータの一般属性特徴を抽出するための機能と、前記一般属性特徴は、
複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ
、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテ
ゴリに属する確度を示し、

10

前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の
確度を決定するための機能であって、

前記一般属性特徴を条件付き確率モデルの所与情報として設定する機能と、

前記条件付き確率モデルを用いて、前記現行クエリが前記所与情報の条件下で前記そ
れぞれのカテゴリに属する複数の事後確率を計算する機能と、

前記複数の事後確率を前記複数の確度として設定する機能とを含み、

前記複数の確度に少なくとも部分的に基づいてカテゴリを選択する機能と、選択された
カテゴリである選択カテゴリは前記複数のカテゴリの１つであり、他のカテゴリの確度よ
りも高い確度を有し、

20

前記選択カテゴリ内で、前記現行クエリデータに対応する検索結果を検索するための機
能と、

前記検索結果を返すための機能と、

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。

【請求項１７】

請求項１に記載の方法であって、前記カテゴリの選択は、前記複数のカテゴリから予め
定められた数のカテゴリを選択することを含む、方法。

【請求項１８】

請求項１に記載の方法であって、前記カテゴリの選択は、前記複数のカテゴリから予め
定められたしきい値を超えた１つのカテゴリを選択することを含む、方法。

30

【発明の詳細な説明】

【技術分野】

【０００１】

〔他の出願の相互参照〕

本願は、すべての目的のために参照により本明細書に組み込まれる、２００９年９月２
７日出願の発明の名称を「SEARCH METHOD AND DEVICE」とする
、中国特許出願第２００９１０１７５０６６．３号に基づく優先権を主張する。

【０００２】

本発明は、情報処理の分野に関し、特に情報検索に関する。

40

【背景技術】

【０００３】

検索エンジンは、一般に、情報を取得するためにインターネットユーザによって利用さ
れる。電子商取引プラットフォームで用いられる検索エンジンの１つのタイプは、カテ
ゴリベース検索エンジンである。かかるプラットフォームでは、ユーザが、１または複数の
カテゴリ（スポーツ、エンターテイメントなど）に対応するクエリキーワードを入力する
。通常のカテゴリベース検索エンジンは、検索キーワードに対応するカテゴリを正確に特
定することができる限りは、正確に検索範囲を狭めて関係のある製品をユーザに薦めるこ
とができる。カテゴリベース検索エンジンの中には、クエリキーワードおよび産業カテ
ゴリのマッピングテーブルなど、ルックアップテーブルを用いるものもある。かかるマッピ

50

ングテーブルの一例を表 1 に示す。

【表 1】

クエリワード	カテゴリ 1	カテゴリ 2	カテゴリ 3
電動バイク	輸送	スポーツおよび エンターテイメント	自動車およびオートバイ
...

【0004】

例えば、ユーザがクエリワード「電動バイク」を入力した場合、検索エンジンは、マッ
ピングテーブル内でクエリワードを検索し、次いで、3つのマッチングカテゴリをクライ
アントに返すので、ユーザは3つのカテゴリの中から1つを選択することが可能であり、
それにより、検索エンジンを通してユーザにとって関心のあるカテゴリ内で検索を行い、
検索エンジンから返された検索結果を受信する。

10

【0005】

カテゴリベースの検索エンジンでは、マッピングテーブル内のクエリデータおよび対応
するカテゴリの両方が手動で設定されるため、処理できるクエリキーワードは、マッピン
グテーブル内でクエリデータが属するカテゴリの数に限定されうる。したがって、検索エ
ンジンは、ユーザによって入力されたクエリワードがマッピングテーブル内のクエリデー
タと一致する場合にのみユーザに検索結果を返すことができる。この結果、従来技術にお
けるこの検索方法では、検索データのカバー率が低くなる。また、マッピングテーブル内
のクエリデータおよび対応するカテゴリの両方が手動で設定されるため、マッピングテー
ブルの維持管理コストが高い傾向にある。さらに、任意の産業カテゴリが修正された場合
に、もともとのマッピングテーブルを更新または再構築する必要があるため、発展性に乏
しい。

20

【図面の簡単な説明】

【0006】

以下の詳細な説明と添付の図面において、本発明の様々な実施形態を開示する。

【0007】

【図 1】検索処理の一実施形態を示すフローチャート。

30

【0008】

【図 2】クエリ語および製品の間の関係を示すユーザインターフェースの一例を示す図。

【0009】

【図 3】カテゴリへのリンクを示すユーザインターフェースの一例を示す図。

【0010】

【図 4】検索処理の別の実施形態を示すフローチャート。

【0011】

【図 5】検索応用例の別の実施形態を示すフローチャート。

【0012】

【図 6】クエリカテゴリライザの入力および出力を示すブロック図。

40

【0013】

【図 7】応用例に従って、検索方法の別の実施形態を示すフローチャート。

【0014】

【図 8】検索システムの一実施形態を示すブロック図。

【0015】

【図 9】検索システムの別の実施形態を示すブロック図。

【0016】

【図 10】検索システムの別の実施形態を示すブロック図。

【0017】

【図 11】検索システムの別の実施形態を示すブロック図。

50

【発明を実施するための形態】

【0018】

本発明は、処理、装置、システム、物質の組成、コンピュータ読み取り可能な格納媒体上に具現化されたコンピュータプログラム製品、および/または、プロセッサ（プロセッサに接続されたメモリに格納および/またはそのメモリによって提供される命令を実行するよう構成されたプロセッサ）を含め、様々な形態で実装されうる。本明細書では、これらの実装または本発明が取りうる任意の他の形態を、技術と呼ぶ。一般に、開示された処理の工程の順序は、本発明の範囲内で変更されてもよい。特に言及しない限り、タスクを実行するよう構成されるものとして記載されたプロセッサまたはメモリなどの構成要素は、ある時間にタスクを実行するよう一時的に構成された一般的な構成要素として、または、タスクを実行するよう製造された特定の構成要素として実装されてよい。本明細書では、「プロセッサ」という用語は、1または複数のデバイス、回路、および/または、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構成された処理コアを指すものとする。

10

【0019】

以下では、本発明の原理を示す図面を参照しつつ、本発明の1または複数の実施形態の詳細な説明を行う。本発明は、かかる実施形態に関連して説明されているが、どの実施形態にも限定されない。本発明の範囲は、特許請求の範囲によってのみ限定されるものであり、多くの代替物、変形物、および、等価物を含む。以下の説明では、本発明の完全な理解を提供するために、多くの具体的な詳細事項が記載されている。これらの詳細事項は、例示を目的としたものであり、本発明は、これらの具体的な詳細事項の一部または全てがなくとも特許請求の範囲に従って実施可能である。簡単のために、本発明に関連する技術分野で周知の技術事項については、本発明が必要以上にわかりにくくならないように、詳細には説明していない。

20

【0020】

クエリデータの一般属性特徴に基づくデータ検索について説明する。一般属性特徴は、検索されるデータ内で利用可能なカテゴリに対する現行クエリデータの確度（*confidence degrees*）を計算するために用いられる。適切なカテゴリが確度に基づいて選択され、クエリの結果を取得するために検索される。

【0021】

図1は、検索処理の一実施形態を示すフローチャートである。処理100は、例えば、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携帯型デバイス、フラットパネルデバイス、マルチプロセッサデバイス、ならびに、それらのいずれかを含む分散コンピュータ環境など、数多くの汎用または専用のコンピュータデバイス環境または構成において実行できる。処理は、コンピュータによって実行されるコンピュータ実行可能な命令の一般的な形態（例えば、プログラムモジュール）内に記述されうる。一般に、プログラムモジュールは、特定のタスクの実行または特定の抽象データ型の実装のためのルーチン、プログラム、オブジェクト、コンポーネント、データ構造などを備える。アプリケーションは、通信ネットワークで接続された複数のリモート処理デバイスによってタスクが実行され、格納デバイスを含むローカルまたはリモートコンピュータ格納媒体にプログラムモジュールが格納されうる分散コンピュータ環境に実装されてもよい。

30

40

【0022】

工程101では、現行クエリデータが受信される。いくつかの実施形態では、現行クエリデータは、ブラウザまたはその他の適切なユーザインターフェースを介して送信され、カテゴリベース検索エンジンサーバによって受信される。

【0023】

工程102では、現行クエリデータの一般属性特徴が抽出される。一般属性特徴は、検索エンジンによって用いられる様々なカテゴリに対する現行クエリデータの確度を計算するために用いられる。

【0024】

50

様々な実施形態において、一般属性特徴は、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、セッション特徴、または、それらの組み合わせを含む。別の実施形態で、他の特徴が用いられてもよい。

【0025】

本明細書で用いられているように、単語レベル特徴とは、現行クエリデータ内のキーワードに対応する。いくつかの実施形態では、非ストップワードに対応する語根が、単語レベル特徴として抽出される。例えば、現行クエリデータが「find Nokia N95 phones」である場合、単語レベル特徴は、「find」、「Nokia」、「N95」、および、「phone」を含む。いくつかの実施形態において、現行クエリデータのキーワードは、周知の単語分割技術を用いて取得できる。

10

【0026】

構文特徴とは、現行クエリデータ内の見出し語に対応する語根を指す。見出し語とは、現行クエリデータの主な意図を反映する中核語(core word)を差す。例えば、「find Nokia N95 phones」というクエリにおいては、「phones」が見出し語である。現行クエリデータ内の見出し語に対応する語根が抽出され、構文特徴として用いられる。上記の例では、語根「phone」が、構文特徴として抽出される。

【0027】

意味特徴とは、キーワードの意味または目的を特定するキーワードに対応する意味タグを指す。いくつかの実施形態において、様々な意味タグは、キーワードを動作動詞として特定するための動作タグ、キーワードを製品として特定するための製品タグ、キーワードをブランドとして特定するためのブランドタグ、キーワードを型式として特定するための型式タグなどを含む。いくつかの実施形態では、動作、製品、ブランド、または、型式のいずれでもないキーワードを特定するために、他のタグが用いられる。例えば、クエリ語「find Nokia N95 phones」については、対応する意味特徴として、「動作：find」、「ブランド：Nokia」、「型式：N95」、および、「製品：phone」が含まれる。

20

【0028】

クリック特徴とは、事前設定されたクリックログ内で現行クエリデータに対応する1または複数の被クリックカテゴリを指しており、クリックログとは、クエリデータとユーザによって選択された実際の1または複数のカテゴリとのマッピングを格納するためのデータベースである。例えば、ユーザが、検索クエリ「mp3」を入力して、検索エンジンから返された結果の中の製品をクリックした時に、クリックされた結果は、「Digital Mp3 Player with Lcd And Fm (IMC278A)」という名前を有し、製品カテゴリ「MP3プレーヤ」に属する製品に関する。検索クエリおよび製品カテゴリのマッピングは、クリックログに格納される。図2は、クエリ語および製品の間の関係を示すユーザインターフェースの一例の図である。図の例において、「mp3」は、ユーザによって入力されたクエリ語であり、ユーザによってクリックされた製品は、カテゴリ「mp3プレーヤ」および「音楽プレーヤ」に属する。したがって、クエリ語「mp3」は、クリック特徴「mp3プレーヤ」および「音楽プレーヤ」を有する。

30

40

【0029】

選択特徴とは、事前設定された選択ログ内で現行クエリデータに対応する1または複数の被選択カテゴリを指す。選択ログは、ユーザが入力したクエリワードとクリックされたカテゴリとの間の対応関係を格納するために用いられる。例えば、ユーザが、URL「www.aliibaba.com」を有する電子商取引ウェブサイトで検索ワード「mp3」を入力した場合、検索された製品へのリンクが、これらの製品に対応するカテゴリと共に格納され、その結果、ユーザは、カテゴリへのリンクをクリックすることによって、検索範囲を決定することができる。図3は、カテゴリへのリンクを示すユーザインターフェースの一例の図である。この例において、検索キーワード「mp3」は、「MP3プレーヤ」、「イヤホンおよびヘッドホン」などのカテゴリに対応する。ユーザは、対応するリ

50

リンクを選択することによってカテゴリ内で以前に選択された製品にアクセスできる。

【0030】

本明細書で用いられているように、特定の期間中に同じウェブサイトで行われる一連のクエリ動作は、セッションに属する。セッション特徴とは、同じセッションにおける現行クエリの前のクエリに対応する最適なカテゴリを指す。すなわち、ウェブページを開いてからウェブページを閉じるまでにユーザによって入力されたクエリキーワードは、同じセッションに属すると見なすことができる。例えば、セッションが、3つの連続的なクエリ「phone」、「Nokia phone」、および、「Nokia n95」を含む場合、「Nokia n95」のセッション特徴は、先行するクエリ「Nokia phone」に対応する最良のカテゴリである。「phone」のセッション特徴は、先行するクエリがないため、空値である。

10

【0031】

図1に戻ると、工程103で、利用可能なカテゴリに対する現行クエリデータの確度が、一般属性特徴に少なくとも部分的に基づいて決定される。

【0032】

いくつかの実施形態では、工程102で抽出された現行クエリデータの各特徴が、条件付き確率モデルの所与条件と見なされる。条件付き確率モデルにおいて、確度は、0から1までの範囲の値を取り、所与条件の下で事象が起きる確率を表す。条件付き確率とは、別の事象（条件）が起きる前提条件下で事象が起きる確率を指すため、現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴はそれぞれ条件と見なされ、カテゴリについて計算された確度は、現行クエリデータがそのカテゴリに属する確率であると考えられる。確度の値が大きくなるほど、現行クエリデータが対応するカテゴリに属する可能性は高くなる。

20

【0033】

いくつかの実施形態では、対応する確度が、最大エントロピー技術を用いて特定のカテゴリに対して計算される。あるスケールを有する訓練クエリデータを利用して、ユーザのクエリをシミュレートし、クエリデータがカテゴリ分けされる確率を決定するために、確度は、時に、いくつかの実施形態において機械学習技術を用いてそれぞれのカテゴリについて計算される。例えば、単語レベル特徴について、「phone」を含むクエリがエレクトロニクスカテゴリに属する」ことを示す記録が訓練データにある場合、現行のクエリデータにおける単語レベル特徴「phone」について、単語レベル特徴は、エレクトロニクスカテゴリに対する確度を計算すると、対応する値「1」を取り、エレクトロニクスカテゴリ以外のカテゴリに対する確度を計算すると「0」の値を取る。同様に、現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を、それぞれ、訓練データの内容に従って計算することが可能であり、これらの特徴値の存在下で、現行クエリデータがカテゴリに属する確率を条件付き確率モデルに従って計算することができる。

30

【0034】

工程104では、カテゴリが、確度に少なくとも部分的に基づいて、利用可能なカテゴリから選択される。カテゴリの選択は、様々な方法で行うことができるいくつかの実施形態では、高い確度を有するいくつかのカテゴリのサブセットが送信される。例えば、40の利用可能なカテゴリが存在し、40の対応する確度が計算されてソートされたと仮定する。検索エンジンサーバは、それぞれのカテゴリおよびそれらに対応する確度をユーザに返すか、あるいは、高い確度を有する上位10のカテゴリをユーザに返し、その結果として、ユーザは、必要に応じて製品の情報をより正確に検索するために、ユーザにとって関心のあるカテゴリに従って、検索範囲を手動で決定することができる。いくつかの実施形態では、サーバは、確度を比較し、最も高い確度を有するカテゴリを自動的に選択する。

40

【0035】

工程105では、選択済みカテゴリは、現行クエリデータを用いて検索され、工程106で、検索結果がクライアントに返される。

50

【 0 0 3 6 】

検索エンジンサーバは、選択済みカテゴリを通知され、現行クエリデータに関する情報を選択済みカテゴリの範囲内で検索する。いくつかの実施形態では、クライアントから送信された現行クエリデータから特徴が抽出され、ここで、特徴は、現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴に関し、これらの特徴は、訓練データ内でマッチングされ、それにより、一致結果を示す特徴値が取得される。現行クエリデータがそれぞれのカテゴリに属することを示す確度が、条件付き確率モデルに従って特徴値から計算され、クエリデータを検索するのに適切なカテゴリを選択するために用いられる。したがって、検索エンジンサーバの動作効率および性能、ならびに、検索結果の精度が改善される。

10

【 0 0 3 7 】

図 4 は、検索処理の別の実施形態を示すフローチャートである。

【 0 0 3 8 】

工程 4 0 1 では、クエリカテゴリライザが、現行クエリデータを受信する。この実施形態において、クエリカテゴリライザは、現行クエリデータに関連する特徴を抽出し、抽出された特徴についてカテゴリに対応する確度を計算するために用いられる。クエリカテゴリライザは、ソフトウェアまたはハードウェアとして実装可能であり、検索エンジンサーバの追加的な機能として、または、検索エンジンサーバに接続された別個のエンティティとして追加できる。

【 0 0 3 9 】

工程 4 0 2 では、クエリカテゴリライザは、現行クエリデータの単語レベル特徴を抽出する。単語レベル特徴を抽出する方法の技術の例については上述した。

20

【 0 0 4 0 】

工程 4 0 3 では、クエリカテゴリライザは、現行クエリデータの構文特徴を抽出する。この実施形態では、まず、現行クエリデータのそれぞれの単語の品詞が、品詞タグ付けツールを用いてタグ付けされる。例えば、「phones」の品詞はnnsであり、単語が複数名詞であることを示す。次に、現行クエリデータの句が、句特定ツールを用いて特定される。例えば、現行クエリデータ「find Nokia N95 phones」は、2つの句、すなわち、それぞれ、動詞句「find」および名詞句「Nokia N95 phones」を含む。次いで、見出し語（「phones」など）が、ルールに従って名詞句から抽出される。例えば、ルールは、「クエリワードが名詞句を1つだけ含む場合、名詞句の最後の名詞が見出し語である」と規定してよい。最後に、「phones」の語根「phone」が、語根抽出ツールを用いて抽出される。標準的または専用の品詞タグ付けツール、句特定ツール、および、語根抽出ツールがいくつかの実施形態で用いられる。

30

【 0 0 4 1 】

工程 4 0 4 では、クエリカテゴリライザは、現行クエリデータの意味特徴を抽出する。この例では、ブランド名（例えば、「Nokia」）のリスト、モデル名（例えば、「N95」）のリスト、および、基本製品名（例えば、「phone」）のリストが、検索エンジンサーバに格納されたクエリログに基づいて電子商取引ネットワーク上で公開された製品情報から生成される。したがって、クエリカテゴリライザは、上述のそれぞれのリストに従って現行クエリデータのキーワードに含まれるブランド、モデル、製品名などにタグ付けし、キーワードの語根を意味特徴として抽出する。

40

【 0 0 4 2 】

工程 4 0 5 では、クエリカテゴリライザは、現行クエリデータのクリック特徴を抽出する。いくつかの実施形態において、クエリカテゴリライザは、現行クエリデータおよびクリック特徴の間の対応関係に従って、事前設定されたクリックログから直接的に、対応するクリック特徴を取得する。いくつかの実装例において、異なる製品が同じカテゴリ内の情報に関連する場合があります。ユーザは、同じクエリワードについて複数の製品に対応するウェブページリンクをクリックしてもよい。そのため、同じクエリワードが、クリックログ内の複数

50

のカテゴリに対応しうる。検索エンジンサーバは、ユーザが入力したクエリワードと、検索エンジンから薦められた製品の中からユーザがクリックする製品のウェブページリンクとを記録して、これらの製品が属する異なるカテゴリを記録し、ひいては、クエリワードおよびカテゴリの間の対応関係を確立し、クエリワード、カテゴリ、および、対応関係に基づいてクリックログを作成する。

【0043】

いくつかの実施形態では、まず、事前設定されたクリックログ内で現行クエリデータに対応する上位 n 個の最頻カテゴリ（いくつかの実施形態では、 n が 3 に設定されることが好ましい）が、クリック特徴としてクリックログから取得される。現行クエリデータがクリックログにおいて一致のない場合、現行クエリデータに最も近いクエリがクリックログ内で検索され、クリックログ内で最も近いクエリに対応する上位 n 個の最頻カテゴリが、現行クエリデータのクリック特徴として決定される。現行クエリデータから最小の編集距離にあるクエリデータが、最も近いクエリとしてクリックログ内で検索される。編集距離とは、2つのクエリデータを同一にするようにそれらに対して挿入、削除、修正などの動作を実行する際の動作のコストを指す。編集コストは、編集距離と正比例する。

10

【0044】

工程 406 では、クエリカテゴリライザは、現行クエリデータの選択特徴を抽出する。いくつかの実施形態において、クエリカテゴリライザは、現行クエリデータおよび選択特徴の間の対応関係に従って、事前設定された選択ログから直接的に、対応する選択特徴を取得する。いくつかの実装例において、検索エンジンは、同じクエリワードが選択ログ内の複数のカテゴリに対応しうるように、マッピングテーブルに従って同じクエリワードに対して複数のカテゴリを推奨してもよい。したがって、検索エンジンサーバは、マッピングテーブルに記録されたカテゴリと、カテゴリおよびクエリワードの間の対応関係とに基づいて選択ログを作成できる。

20

【0045】

いくつかの実施形態では、まず、事前設定された選択ログ内でクエリに対応する上位 n 個の最頻カテゴリ（ n は 3 であることが好ましい）が、選択特徴として選択ログから取得される。現行クエリデータが選択ログにない場合、現行クエリデータに最も近いクエリが選択ログ内で検索され、選択ログ内で最も近いクエリに対応する上位 n 個の最頻カテゴリが、クエリの選択特徴として決定される。現行クエリデータから最小の編集距離のクエリデータが、最も近いクエリとして検索される。

30

【0046】

工程 407 では、クエリカテゴリライザは、現行クエリデータのセッション特徴を抽出する。電子商取引の分野では、セッションは、通例、同じ対象物に関して実行される。同じセッション内のクエリデータが属するカテゴリは、通例、同一のカテゴリまたは関連したカテゴリである。実際には、例えば、セッションは、3つの連続的なクエリ、すなわち、「phone」、「Nokia phone」、および、「Nokia n95」を含み、それらはすべてユーザのクエリ意図を反映しており、3つのクエリが属するカテゴリは関連している。したがって、セッション特徴も、現行クエリデータがカテゴリに属する確度を或る程度反映しうる。

40

【0047】

工程 408 では、クエリカテゴリライザは、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を、条件付き確率モデルの所与情報として設定する。

【0048】

任意の与えられたカテゴリ C について、現行クエリデータがカテゴリ C に対応する確度は、現行クエリデータが所与の関連する情報 x の条件下でカテゴリ C に属すると決定される事後確率 p （「 c が正確なカテゴリ | x 」と表現される）として表すことができる。ここで、 x は、所与情報、すなわち、この実施形態では、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を表す。

50

【 0 0 4 9 】

工程 4 0 9 では、クエリカテゴライザは、条件付き確率モデルを用いて、現行クエリが所与情報の条件下でそれぞれのカテゴリに属する事後確率を計算し、現行クエリデータがそれぞれのカテゴリに属する確度として事後確率を設定する。

【 0 0 5 0 】

条件付き確率モデルは、現行クエリデータによって取得された特徴の加重を訓練することができ、現行クエリデータがそれぞれのカテゴリに属すると決定される条件付き確率（すなわち、確度）は、条件付き確率モデルを用いて、取得済みの対応する特徴値および加重から計算される。

【 0 0 5 1 】

以下の式を用いて規定される最大エントロピモデルが、いくつかの実施形態で条件付き確率モデルとして用いられる：

【 数 1 】

$$p(y|x) = \frac{1}{Z(x)} \exp \left(\sum_j \lambda_j f_j(x, y) \right) \quad (1)$$

【 0 0 5 2 】

ここで、 $y = \{c \text{ が正確なカテゴリ、} c \text{ が不正確なカテゴリ}\}$ 、 x は現行クエリデータに対応する所与情報、すなわち、この実施形態では単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴である。 f_j は、最大エントロピモデルに対応する特徴値であり、例えば、 x が単語レベル特徴である時に、訓練データが「現行クエリデータのキーワードが「phone」を含めば、カテゴリはCである」ことを示す場合、特徴 f_j は、「現行のクエリデータのキーワードが「phone」を含む」という条件で値「1」を有し、そうでなければ、 f_j は値「0」を有する。 λ_j は、現行クエリデータの特徴に対応する加重であり、モデル訓練を通して得ることができる。 $Z(x)$ は、モデル訓練を通して得ることができる規格化因子である。

【 0 0 5 3 】

実際的には、条件付き確率モデルは、線形回帰など機械学習技術を用いて生成できる。いくつかの実施形態において、サポートベクターマシンモデルが、確度として用いるスコアを計算するために採用されてもよいが、サポートベクターマシンモデルは、一部では条件付き確率モデルと見なされていない。

【 0 0 5 4 】

工程 4 0 1 0 では、クエリカテゴライザは、計算および取得されたそれぞれの確度が、事前設定された第 1 の閾値より高いか否かを判定し、高い場合にはフローは工程 4 0 1 2 に進み、高くない場合にはフローは工程 4 0 1 1 に進む。

【 0 0 5 5 】

現行のクエリデータがそれぞれのカテゴリに属する確度を計算して取得した後、クエリカテゴライザは、第 1 の閾値を事前設定することができる。例えば、閾値が 0.5 に設定された場合、対応する確度が 0.5 よりも高いカテゴリが、確度の高いカテゴリと見なされる。

【 0 0 5 6 】

工程 4 0 1 1 では、確度の内、第 1 の閾値より低い確度が記録される。それぞれのカテゴリについて計算および取得された確度がすべて第 1 の閾値より低い場合、所定の数の相対的に高い確度がすべての確度から直接取得され、それらに対応するカテゴリが取得され、フローは工程 4 0 1 3 に進む。

【 0 0 5 7 】

すべての確度の判定が終わった時に、それぞれのカテゴリに対応する確度に第 1 の閾値を超えるものがない場合、クエリカテゴライザは、すべての確度を降順で直接ソートして、第 1 の所定の数の比較的高い確度およびそれらに対応するカテゴリを取得する。

【 0 0 5 8 】

工程 4 0 1 2 では、クエリカテゴライザは、閾値を超える確度と、それらに対応するカテゴリとを取得する。

【 0 0 5 9 】

工程 4 0 1 3 では、クエリカテゴライザは、閾値を超える確度およびそれらに対応するカテゴリ、もしくは、所定の数の確度およびそれらに対応するカテゴリを、ユーザに返す。

【 0 0 6 0 】

いくつかの実施形態において、クエリカテゴライザは、第 1 の閾値を超えるすべての確度およびそれらに対応するカテゴリをクライアントに返し、さらに、確度を降順で提示する；あるいは、クエリカテゴライザは、第 1 の所定の数（例えば、8）の確度およびそれらに対応するカテゴリをクライアントに提示する。

10

【 0 0 6 1 】

工程 4 0 1 4 では、クエリカテゴライザは、確度に従ってクライアントが選択した 1 つのカテゴリ内で現行クエリデータを検索する。

【 0 0 6 2 】

この実施形態では、標準的な単語分割ユニット、品詞タグ付けツールなど、既存のツールを用いることができる。現行クエリデータがこれまでに出現したか否かにかかわらず条件付き確率モデルを用いて様々な抽出済みの特徴から、それぞれ可能性のある関連のカテゴリの確度を取得できるため、条件付き確率に基づいた機械学習方法で高いカバー率を得ることが可能であり、したがって、幅広い用途および高い精度を達成できる。

20

【 0 0 6 3 】

図 5 は、検索応用例の別の実施形態を示すフローチャートである。

【 0 0 6 4 】

工程 5 0 1 では、クエリカテゴライザは、現行クエリデータ内のキーワードを取得する。この実施形態のクエリカテゴライザは、現行クエリデータ内のキーワードを取得するための単語分割ユニットと統合されている。例えば、現行クエリデータが「find Nokia N95 phones」であるとする、カテゴリ C について、「現行のクエリワードがカテゴリ C に対応し「find」を含む」、「現行のクエリワードがカテゴリ C に対応し「Nokia」を含む」、および、「現行のクエリワードがカテゴリ C に対応し「N95」を含む」という条件下で、確度の関数 $f(x, y)$ が「1」であり、そうでない条件下で、 $f(x, y)$ が「0」である場合、 $f(x = \text{find}, \text{「find Nokia N95 phones」カテゴリ C})$ は、「find」という単語レベル特徴が「find Nokia N95 phones」から抽出された時には、値「1」を有する。同様に、 $f(x = \text{Nokia}, \text{「find Nokia N95 phones」カテゴリ C})$ および $f(x = \text{N95}, \text{「find Nokia N95 phones」カテゴリ C})$ も値「1」を有するが、 $f(x = \text{phone}, \text{「find Nokia N95 phones」カテゴリ C})$ は値「0」を有する。

30

【 0 0 6 5 】

工程 5 0 2 では、クエリカテゴライザは、現行クエリデータ内のキーワードの品詞にタグ付けし、現行クエリデータ内のキーワードの句を特定する。

40

【 0 0 6 6 】

この実施形態のクエリカテゴライザは、品詞タグ付けツール、句特定ツール、および、語根抽出ツールと統合される。各キーワードの品詞が、品詞タグ付けツールを用いてタグ付けされ、例えば、現行クエリデータが「find Nokia N95 phones」である場合、キーワード「phones」の品詞は、単語が複数名詞であることを示す nns であり、現行クエリデータは、動詞句「find」および名詞句「Nokia N95 phones」を含む。

【 0 0 6 7 】

工程 5 0 3 では、クエリカテゴライザは、上述のタグ付けおよび特定の結果から現行クエリデータ内の名詞句を取得し、名詞句の見出し語を抽出する。この例では、クエリカテ

50

ゴライザは、名詞句「Nokia N95 phones」から見出し語「phones」を抽出する。

【0068】

工程504では、クエリカテゴライザは、語根抽出ツールを用いて見出し語の語根を抽出する。この例では、「phones」の語根「phone」が、語根抽出ツールを用いて抽出される。

【0069】

この実施形態で記載された工程502～504は、現行クエリデータの構文特徴を抽出するために用いられる。「find Nokia N95 phones」については、構文特徴が「phone」である時に、カテゴリCに属するデータが訓練データ内でマッチングされうる場合、「find Nokia N95 phones」の構文特徴は、値「1」を有し、そうでない場合、値「0」を有する。現行クエリワードは、カテゴリCに対応し、クエリワードの見出し語は、語根「phone」を含む。したがって、 $f(x = \text{phone}, \text{find Nokia N95 phones})$ カテゴリC)は、カテゴリCの確度を計算するために値「1」を取り、そうでない場合には値「0」を取る。

【0070】

工程505では、クエリカテゴライザは、現行クエリデータのキーワードの種類を特定する。すなわち、現行クエリデータのそれぞれのキーワードは、種類でタグ付けされ、例えば、「find」は動作動詞としてタグ付けされ、「Nokia」はブランド名としてタグ付けされる、などである。

【0071】

工程506では、クエリカテゴライザは、種類の特定に従って、キーワードの動作動詞タグ、製品タグ、ブランドタグ、型式タグ、および/または、別のタグを取得する。クエリ「find Nokia N95 phones」については、全部で4つの意味特徴、すなわち、「動作：find」、「ブランド：Nokia」、「型式：N95」、および、「製品：phone」が取得される。

【0072】

工程505～506は、現行クエリデータの意味特徴を取得するために用いられる。

【0073】

工程507では、クエリカテゴライザは、現行クエリデータが、事前設定されたクリックログ内に存在するか否かを判定する。ここで、事前設定されたクリックログは、クエリデータおよびクリック特徴の間の対応関係を格納するよう適合されている。存在する場合、フローは工程508に進み、存在しない場合、フローは工程509に進む。

【0074】

上述のように、検索エンジンサーバは、ユーザが入力した現行クエリデータに回答して、現行クエリデータに関連した製品への複数のリンクを返すことができる。特に、クリックログは、ユーザがクエリワードを入力した後にクリックする製品へのリンクと、製品が属するカテゴリとを記録する。特に、クリックログは、さらに、製品タイトル、製品キーワードなどに関する情報を含むことができる。

【0075】

工程508では、現行クエリデータに対応する第1の所定の数の被クリックカテゴリが、事前設定されたクリックログから取得される。ここで、被クリックカテゴリとは、クエリデータの検索後にユーザによってクリックされた対応する製品が属するカテゴリである。

【0076】

例えば、現行クエリデータが「find Nokia N95 phones」である場合、対応するクリック特徴は、「クリック__カテゴリ__トップ1：遠隔通信」、「クリック__カテゴリ__トップ2：家庭用電化製品」、および、「クリック__カテゴリ__トップ3：コンピュータハードウェアおよびソフトウェア」を含む。ここで、トップ*i* (*i* = 1, 2, 3)は、クリック特徴が属する上位*i*番目の被クリックカテゴリを特定し、クリッ

10

20

30

40

50

ク__カテゴリは、被クリックカテゴリを表す。

【 0 0 7 7 】

工程 5 0 9 では、クエリカテゴリライザは、現行クエリデータに最も近い被クリッククエリデータについてクリックログに照会する。ここで、被クリッククエリデータは、現行クエリデータから最短編集距離にあるクリックログ内のデータである。クエリカテゴリライザは、さらに、被クリッククエリデータに対応する第 1 の所定の数のカテゴリを現行クエリデータのクリック特徴として取得する。

【 0 0 7 8 】

工程 5 0 7 ~ 5 0 9 は、現行クエリデータのクリック特徴を抽出するために用いられる。

10

【 0 0 7 9 】

工程 5 0 1 0 では、クエリカテゴリライザは、現行クエリデータが、事前設定された選択ログ内に存在するか否かを決定する。ここで、選択ログは、クエリデータおよび選択特徴の間の対応関係を格納するよう適合されている。存在する場合、フローは工程 5 0 1 1 に進み、存在しない場合、フローは工程 5 0 1 2 に進む。

【 0 0 8 0 】

工程 5 0 1 1 では、クエリカテゴリライザは、事前設定された選択ログから現行クエリデータに対応する第 1 の所定の数の被選択カテゴリを取得する。ここで、被選択カテゴリとは、検索エンジンのマッピングテーブル内で現行クエリデータと関連付けられたカテゴリを表す。

20

【 0 0 8 1 】

工程 5 0 1 2 では、クエリカテゴリライザは、現行クエリデータに最も近い被選択クエリデータについて選択ログに照会し（ここで、被選択クエリデータは、現行クエリデータから最短編集距離にある選択ログ内のデータである）、被選択クエリデータに対応する第 1 の所定の数のカテゴリを現行クエリデータの選択特徴として取得する。

【 0 0 8 2 】

工程 5 0 1 0 ~ 5 0 1 2 は、現行クエリデータの選択特徴を抽出するために用いられる。

【 0 0 8 3 】

工程 5 0 1 3 では、クエリカテゴリライザは、同じセッションにおける現行クエリデータの前のクエリデータを取得する。この例では、クエリカテゴリライザは、「 f i n d N o k i a N 9 5 p h o n e s 」の前のクエリデータについて照会する。

30

【 0 0 8 4 】

工程 5 0 1 4 では、クエリカテゴリライザは、前のクエリデータに対応する最適なカテゴリ分け結果が属するカテゴリを取得し、最適なカテゴリ分け結果が属するカテゴリを現行クエリデータのセッション特徴として設定する。例えば、ユーザが 3 つのクエリ q_1 、 q_2 、および、 q_3 を入力したと仮定する。クエリカテゴリライザは、 q_1 に対応する最適カテゴリ c_1 を決定する。クエリカテゴリライザが q_2 を処理する時、対応するセッション特徴は、前のクエリデータに対応する最適なカテゴリ分け結果がカテゴリ c_1 に属することである。

40

【 0 0 8 5 】

工程 5 0 1 5 では、クエリカテゴリライザは、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴に従って、それぞれのカテゴリに対応する現行クエリデータの確度を計算して取得する。ここで、確度は各々、現行クエリデータがカテゴリの 1 つに属する確度を指す。

【 0 0 8 6 】

具体的な例では、この実施形態において、現行クエリデータの単語レベル特徴がそれぞれ、値 1、1、1、および、0 を有し、構文特徴がそれぞれ、値 1 および 0 を有し、意味特徴がそれぞれ、値 1、1、1、0、および、0 を有し、クリック特徴がそれぞれ、値 0、0、および、1 を有し、選択特徴がそれぞれ、値 1、0、および、1 を有し、セッショ

50

ン特徴が値 1 を有すると仮定する。それぞれのカテゴリに対応する確度の値は、以下の最大エントロピモデルの式で得ることができる：

【数 2】

$$p(y|x) = \frac{1}{Z(x)} \exp \left(\sum_j \lambda_j f_j(x, y) \right)$$

【0087】

工程 5016 では、クエリカテゴリザは、それぞれのカテゴリおよびそれらに対応する確度をクライアントに返す。

【0088】

クエリカテゴリザは、確度を計算した後に、対応するカテゴリおよび確度の対応関係を確度の降順でソートし、それらをクライアントに返す。

【0089】

工程 5017 では、検索エンジンサーバは、確度に従ってクライアントが選択した 1 つのカテゴリ内で現行クエリデータを検索する。

【0090】

上記の例からわかるように、検索カテゴリザは、検索エンジンサーバと別個に設けられてもよいし、実際の応用例における検索システムがクエリをカテゴリ分けする場合に、検索エンジンサーバに統合されてもよい。図 6 は、クエリカテゴリザの入力および出力を示すブロック図である。

【0091】

図 7 は、応用例に従って、検索方法の別の実施形態を示すフローチャートである。

【0092】

工程 701 では、クライアントから送信された現行クエリデータが受信される。

【0093】

工程 702 では、現行クエリデータの一般属性特徴が、現行クエリデータの確度を計算するために抽出される。

【0094】

工程 703 では、それぞれのカテゴリに対応する現行クエリデータの確度が、一般属性特徴に従って計算されて取得される。ここで、確度は各々、現行クエリデータがカテゴリの 1 つに属する確度を指す。

【0095】

この実施形態において、工程 701 ~ 703 は、上述の対応する動作と同様であってよい。

【0096】

工程 704 では、現行クエリデータの検索が、利用可能なカテゴリの中で最大の確度を持つと算出されたカテゴリで実行される。

【0097】

この実施形態では、検索エンジンサーバは、最大の確度を持つカテゴリを自動的に選択し、次いで、現行クエリデータをそのカテゴリ内で直接検索する。検索結果が、ユーザに返される。ここで、検索エンジンサーバは、最大の確度に従って検索を実行するだけでなく、特定の条件を満たす特定の確度に対応するカテゴリまたは複数の確度に対応する複数のカテゴリを、現行クエリデータの検索範囲として選択する。なお、当業者がこの解決法に対してないうるいくつかの変形例も本発明の範囲に含まれる。

【0098】

この実施形態に従った方法は、確度およびそれらに対応するカテゴリをユーザに返す代わりに、最適な結果に直接従った検索を実行できるため、一般に、対応するマッピングテーブルを格納することなしに、検索結果に対するユーザの要求を満たすことができる。ユーザの現行クエリデータが過去に出現したか否かに関わらず、現行クエリデータがカテゴリに属する最大確度を決定することが可能であり、最も合理的な検索動作と考えられるカ

10

20

30

40

50

カテゴリで検索を実行するため、結果として、カテゴリおよび確度のデータを検索エンジンからユーザに送信することなしに、クエリデータのカバー率を改善することができる。したがって、帯域幅が節約され、サーバの通信負荷が軽減される。

【0099】

図8は、検索システムの一実施形態を示すブロック図である。システム800は、以下のような1または複数のまたは複数の計算デバイスを用いて実装されてよい。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携帯型デバイス、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に基づくシステム、セットトップボックス、プログラム可能な消費者電子デバイス、ネットワークPC、ミニコンピュータ、大規模コンピュータ、特定目的デバイス、任意の前述のシステムまたはデバイスを含む汎用計算環境、または、1または複数のプロセッサ、および、プロセッサと接続したメモリを備え、指示を有するプロセッサを提供するよう構成された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。

10

【0100】

受信モジュール801は、クライアントから送信された現行クエリデータを受信するよう適合される。

【0101】

抽出モジュール802は、現行クエリデータの一般属性特徴を抽出するよう適合されており、一般属性特徴は、それぞれのカテゴリに対応する現行クエリデータの確度を計算するために用いられる。

20

【0102】

本願の一実施形態では、現行クエリデータのいくつかの特徴が、後の条件付き確率の動作のために抽出され、一般属性特徴は、現行クエリデータから抽出される必要があり、上述のように、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を含むことが好ましい。

【0103】

計算・取得モジュール803は、一般属性特徴に従って、それぞれのカテゴリに対応する現行クエリデータの確度を計算および取得するよう適合されており、ここで、確度は各々、現行クエリデータがカテゴリの1つに属する確度を指す。

【0104】

30

特に、現行クエリデータの抽出された特徴は各々、条件付き確率モデルにおける所与条件とみなすことができ、確度は、事象が所与条件下で起きる確率を表すために0から1までの範囲の値を取る。条件付き確率とは、別の事象(条件)が起きる前提条件下で事象が起きる確率を指すため、この実施形態における現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴はそれぞれ条件と見なされ、カテゴリについて計算された確度は、現行クエリデータがそのカテゴリに属する確率であると考えられうる。したがって、確度が大きくなるほど、現行のクエリデータが対応するカテゴリに属する可能性は高くなる。

【0105】

対応する確度が、最大エントロピーモデルを用いて特定のカテゴリに対して計算されうることが好ましい。かかる技術は、或るスケールを有する訓練クエリデータを利用して、ユーザのクエリ行動をシミュレートし、クエリデータが或るカテゴリにカテゴリ分けされる確率を計算して取得することができるため、確度は、いくつかの実施形態において、機械学習技術を用いてそれぞれのカテゴリについて計算される。これらの特徴の値があれば、条件付き確率モデルに従って、現行クエリデータがカテゴリに属する確率を計算することができる。

40

【0106】

送信モジュール804は、それぞれのカテゴリおよび対応する確度をクライアントに返すよう適合される。

【0107】

50

それぞれ確度が計算される40のカテゴリが実際に存在すると仮定すれば、実際に40の対応する確度が計算され、検索エンジンサーバは、それぞれのカテゴリおよび対応する確度をユーザに返す。その結果として、ユーザは、必要に応じて製品の情報をより正確に検索するために、ユーザにとって関心のあるカテゴリに従って、検索範囲を決定することができる。

【0108】

第1の検索モジュール805は、確度に従ってクライアントが選択した1つのカテゴリ内で、現行クエリデータを検索するよう適合される。

【0109】

この実施形態に記載されたデバイスは、検索エンジンサーバに統合されてもよいし、別個のエンティティとして検索エンジンサーバと接続されてもよい。本願に従った方法は、ソフトウェアとして実装される場合、検索エンジンサーバに追加された新たな機能または対応するプログラム内に別個に書き込まれた新たな機能として実装されてよく、本願は、方法またはデバイスの実装法に関して限定されないことに注意されたい。

【0110】

この実施形態において、デバイスは、確度およびカテゴリの間の対応関係をクライアントに返す場合には、クライアントによって選択された1つのカテゴリ内で現行クエリデータを検索することができる。したがって、検索エンジンサーバの作業負荷をある程度低減し、検索エンジンサーバの動作効率および性能を向上させ、データクエリについてのユーザの検索体験を改善することが可能である。

【0111】

図9は、検索システムの別の実施形態を示すブロック図である。

【0112】

この例において、システム900は、クライアントから送信された現行クエリデータを受信するよう適合された受信モジュール801と、現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を抽出するよう適合された抽出モジュール802と、を備える。

【0113】

所与情報設定サブモジュール901は、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を、条件付き確率モデルの所与情報として設定するよう適合される。

【0114】

確度計算サブモジュール902は、条件付き確率モデルを用いて、現行クエリが所与情報の条件下でそれぞれのカテゴリに属する事後確率を計算し、現行クエリデータがそれぞれのカテゴリに属する確度として事後確率を設定するよう適合される。

【0115】

条件付き確率モデルは、現行クエリデータによって取得された特徴の加重を訓練することができ、現行クエリデータがそれぞれのカテゴリに属すると決定される条件付き確率（すなわち、確度）は、条件付き確率モデルを用いて、取得済みの対応する特徴値および加重から計算される。いくつかの実施形態では、式(1)と関連して上述したような最大エントロピーモデルが用いられる。他の実施形態では、条件付き確率モデルは、線形回帰、サポートベクターマシンモデルなどから生成できる。

【0116】

第1の判定モジュール903は、計算および取得されたそれぞれの確度が、事前設定された第1の閾値より高いか否かを判定するよう適合される。

【0117】

現行クエリデータがそれぞれのカテゴリに属する確度を計算および取得した後、クエリカテゴリは、第1の閾値を事前設定することが可能であり、対応する確度が閾値よりも高いカテゴリは、確度の高いカテゴリと見なされうる。

【0118】

10

20

30

40

50

第1の取得モジュール904は、第1の判定モジュールからの結果がポジティブである時に、第1の閾値を超える確度およびそれらに対応するカテゴリを取得するよう適合される。

【0119】

クエリカテゴリザは、第1の閾値を超えるすべての確度およびそれらに対応するカテゴリをクライアントに返し、確度を降順で提示することができる；あるいは、クエリカテゴリザは、第1の所定の数（例えば、8）の確度およびそれらに対応するカテゴリをクライアントに提示することができる。

【0120】

第2の取得モジュール905は、それぞれのカテゴリについて計算および取得された確度がすべて第1の閾値より低い場合、すべての確度の内の所定の数の相対的に高い確度、および、それらに対応するカテゴリを取得するよう適合される。

10

【0121】

すべての確度の判定が終わった時に、それぞれのカテゴリに対応する確度に第1の閾値を超えるものがない場合、クエリカテゴリザは、すべての確度を降順で直接ソートして、第1の所定の数の比較的高い確度およびそれらに対応するカテゴリを取得する。

【0122】

送信モジュール804は、閾値を超える確度およびそれらに対応するカテゴリ、もしくは、それぞれのカテゴリおよびそれらの確度を、クライアントへ返すよう適合される。

【0123】

20

第1の検索モジュール805は、確度に従ってクライアントが選択した1つのカテゴリ内で、現行クエリデータを検索するよう適合される。

【0124】

図10は、検索システムの別の実施形態を示すブロック図である。

【0125】

システム1000では、受信モジュール801は、クライアントから送信された現行クエリデータを受信するよう適合される。タグ付け・特定サブモジュール1001は、現行クエリデータ内のキーワードの品詞をタグ付けし、現行クエリデータ内のキーワードの句を特定するよう適合される。特に、品詞は品詞タグ付けツールによってタグ付けされ、句は句特定ツールによって特定されることができる。

30

【0126】

見だし抽出サブモジュール1002は、上述のタグ付けおよび特定の結果から現行クエリデータ内の名詞句を取得し、名詞句の見出し語を抽出するよう適合される。

【0127】

語根抽出ツール1003は、見出し語の語根を抽出するよう適合される。

【0128】

種類特定サブモジュール1004は、現行クエリデータのキーワードの種類を特定するよう適合される。

【0129】

タグ取得サブモジュール1005は、種類の特定に従って、キーワードの動作動詞タグ、製品タグ、ブランドタグ、型式タグ、および/または、別のタグを取得するよう適合される。

40

【0130】

第1の判定サブモジュール1006は、現行クエリデータが、事前設定されたクリックログ内に存在するか否かを判定するよう適合されており、ここで、事前設定されたクリックログは、クエリデータおよびクリック特徴の間の対応関係を格納するよう適合されている。

【0131】

第1の取得サブモジュール1007は、第1の判定サブモジュールの結果がポジティブである場合に、事前設定されたクリックログから現行クエリデータに対応する第1の所定

50

の数の被クリックカテゴリを取得するよう適合されており、ここで、被クリックカテゴリとは、クエリデータの検索後にユーザによってクリックされた対応する製品が属するカテゴリである。

【0132】

第1の照会・取得サブモジュール1008は、第1の判定サブモジュールの結果がネガティブである場合に、現行クエリデータに最も近い被クリッククエリデータについてクリックログに照会し（ここで、被クリッククエリデータは、現行クエリデータから最短編集距離にあるクリックログ内のデータである）、被クリッククエリデータに対応する第1の所定の数のカテゴリを現行クエリデータのクリック特徴として取得するよう適合される。

【0133】

第2の判定サブモジュール1009は、現行クエリデータが、事前設定された選択ログ内に存在するか否かを判定するよう適合されており、ここで、選択ログは、クエリデータおよび選択特徴の間の対応関係を格納するよう適合されている。

【0134】

第2の取得サブモジュール1010は、第2の判定サブモジュールの結果がポジティブである場合に、事前設定された選択ログから現行クエリデータに対応する第1の所定の数の被選択カテゴリを取得するよう適合されており、ここで、被選択カテゴリは、検索エンジンのマッピングテーブル内で現行クエリデータと関連付けられたカテゴリを表す。

【0135】

第2の照会・取得サブモジュール1011は、第2の判定サブモジュールの結果がネガティブである場合に、現行クエリデータに最も近い被選択クエリデータについて選択ログに照会し（ここで、被選択クエリデータは、現行クエリデータから最短編集距離にある選択ログ内のデータである）、被選択クエリデータに対応する第1の所定の数のカテゴリを現行クエリデータの選択特徴として取得するよう適合される。

【0136】

第3の取得サブモジュール1012は、同じセッションにおける現行クエリデータの前のクエリデータを取得するよう適合される。

【0137】

第4の取得サブモジュール1013は、前のクエリデータに対応する最適なカテゴリ分け結果が属するカテゴリを取得し、最適なカテゴリ分け結果が属するカテゴリを現行クエリデータのセッション特徴として設定するよう適合される。

【0138】

計算・取得モジュール803は、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴に従って、それぞれのカテゴリに対応する現行クエリデータの確度を計算して取得するよう適合されており、ここで、確度は各々、現行クエリデータがカテゴリの1つに属する確度を指す。

【0139】

送信モジュール804は、それぞれのカテゴリおよびそれらに対応する確度をクライアントに返すよう適合される。

【0140】

第1の検索モジュール805は、確度に従ってクライアントが選択した1つのカテゴリ内で、現行クエリデータを検索するよう適合される。

【0141】

この実施形態において、実装コストを削減するために、単語分割ユニット、品詞タグ付けツールなど、既存のツールを用いることが可能であり、現行クエリデータがこれまでに出現したか否かにかかわらず条件付き確率モデルを用いて様々な抽出済みの特徴から、すべての可能性のある関連カテゴリの確度を取得できるため、条件付き確率に基づいた機械学習方法で高いカバー率を得ることが可能であり、したがって、従来技術のテーブルよりも、幅広い用途および高い精度を達成できる。

【0142】

10

20

30

40

50

図 11 は、検索システムの別の実施形態を示すブロック図である。

【0143】

システム 1100 において、受信モジュール 801 は、クライアントから送信された現行クエリデータを受信するよう適合される。

【0144】

抽出モジュール 802 は、現行クエリデータの確度を計算するために現行クエリデータの一般属性特徴を抽出するよう適合される。

【0145】

計算・取得モジュール 803 は、一般属性特徴に従って、それぞれのカテゴリに対応する現行クエリデータの確度を計算および取得するよう適合されており、ここで、確度は各々、現行クエリデータがカテゴリの 1 つに属する確度を指す。

10

【0146】

第 2 の検索モジュール 1101 は、計算された確度の内の最大確度に対応する 1 つのカテゴリ内で現行クエリデータを検索するよう適合される。

【0147】

この実施形態において、デバイスの第 1 の実施形態との差異は、現行クエリデータがそれぞれのカテゴリに属する確度を計算・取得モジュール 803 が計算して取得した後に、検索エンジンサーバが決定を行い（すなわち、すべての計算された確度の内の最大確度に対応する 1 つのカテゴリを決定し）、その後、そのカテゴリ内で現行クエリデータを直接検索できる点である。さらに、検索結果が直接ユーザに返される。

20

【0148】

この実施形態に従ったデバイスは、確度およびそれらに対応するカテゴリをユーザに返す代わりに、最適な結果に直接従った検索を実行できるため、一般に、カテゴリの情報を記録するためにバックグラウンドで、対応するマッピングテーブルの格納などを行うことなしに、検索結果に対するユーザの要求を満たすことができる。ユーザの現行クエリデータがかつて出現したか否かに関わらず、現行クエリデータがカテゴリに属する最大確度を決定して、最も合理的な検索動作としてそのカテゴリ内で検索を行うことができるため、クエリデータのカバー率を改善し、検索エンジンサーバの性能を最適化し、ユーザの様々な検索要求を満たすことができる。

【0149】

30

検索方法およびデバイスの上述の実施形態に対応して、本願は、さらに、検索システムの一実施形態を提供し、この実施形態において、システムは、特に以下を含む。

【0150】

検索エンジンサーバ側において、

【0151】

クエリカテゴリライザは、受信モジュール、抽出モジュール、計算・取得モジュール、および、送信モジュールを備える。

【0152】

受信モジュールは、クライアントから送信された現行クエリデータを受信するよう適合される。

40

【0153】

抽出モジュールは、現行クエリデータの単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴を抽出するよう適合される。

【0154】

計算・取得モジュールは、単語レベル特徴、構文特徴、意味特徴、クリック特徴、選択特徴、および、セッション特徴に従って、それぞれのカテゴリに対応する現行クエリデータの確度を計算して取得するよう適合されており、ここで、確度は各々、現行クエリデータがカテゴリの 1 つに属する確度を指す。

【0155】

送信モジュールは、それぞれのカテゴリおよびそれらに対応する確度をクライアントに

50

返すよう適合される。

【0156】

第1の検索モジュールは、確度に従ってクライアントが選択した1つのカテゴリ内で現行クエリデータを検索するよう適合される。

【0157】

システムは、さらに、ユーザが入力した現行クエリデータをブラウザを通して受信し、検索エンジンサーバに送信するよう適合されたクライアントを備えてよいことに注意されたい。

【0158】

実施形態に記載したモジュールは、1または複数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、特定の機能を実行するよう設計されたプログラム可能論理デバイスおよび/または特定用途向け集積回路などのハードウェアとして、もしくは、それらの組み合わせとして実装することができる。いくつかの実施形態において、モジュールは、コンピュータデバイス（パーソナルコンピュータ、サーバ、ネットワーク装置など）に本発明の実施形態に記載された方法を実行させるための複数の命令など、不揮発性記憶媒体（光学ディスク、フラッシュ記憶装置、携帯用ハードディスクなど）に格納することができるソフトウェア製品の形態で具現化されてよい。モジュールは、単一のデバイス上に実装されてもよいし、複数のデバイスにわたって分散されてもよい。モジュールの機能は、互いに統合されてもよいし、複数のサブモジュールにさらに分割されてもよい。

【0159】

本明細書ではそれぞれの実施形態が段階的に説明され、各実施形態において他の実施形態との差異が強調されており、共通または同様の要素については相互参照できることに注意されたい。デバイスの実施形態については、方法の実施形態と実質的に同様であるため簡潔に記載されており、その詳細については、方法の実施形態の記載の該当箇所を参照できる。

【0160】

上述の実施形態は、理解しやすいようにいくぶん詳しく説明されているが、本発明は、提供された詳細事項に限定されるものではない。本発明を実施する多くの代替方法が存在する。開示された実施形態は、例示であり、限定を意図するものではない。

適用例1：情報を検索するための方法であって、クライアントから現行クエリデータを受信し、前記現行クエリデータの一般属性特徴を抽出し、前記一般属性特徴は、複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテゴリに属する確度を示し、前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の確度を決定し、選択カテゴリ内において、前記現行クエリデータに対応する検索結果を検索し、前記選択カテゴリは、前記複数のカテゴリの1つであり、前記複数の確度に少なくとも部分的に基づいて選択され、前記検索結果を返すこと、を備える、方法。

適用例2：適用例1に記載の方法であって、前記一般属性特徴は、前記現行クエリデータ内のキーワードに対応する単語レベル特徴を含む、方法。

適用例3：適用例1に記載の方法であって、前記一般属性特徴は、前記現行クエリデータ内の見出し語に対応する語根に対応する構文特徴を含む、方法。

適用例4：適用例3に記載の方法であって、前記構文特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータ内の前記キーワードの品詞タグ付けと、前記現行クエリデータ内の前記キーワードを含む句の特定とを実行し、前記品詞タグ付けおよび前記句の前記特定の結果に少なくとも部分的に基づいて、前記クエリデータ内の名詞句を取得し、前記名詞句の前記見出し語に関連する前記語根を抽出すること、を備える、方法。

適用例5：適用例1に記載の方法であって、前記一般属性特徴は、前記キーワードに対応する意味タグに対応する意味特徴を含む、方法。

適用例6：適用例5に記載の方法であって、前記意味特徴を含む前記一般属性特徴の抽

10

20

30

40

50

出は、前記現行クエリデータの前記キーワードの種類を特定し、前記特定された種類に従って、前記キーワードの動作動詞タグ、製品タグ、ブランドタグ、型式タグ、または、別のタグを取得すること、を備える、方法。

適用例 7：適用例 1 に記載の方法であって、前記一般属性特徴は、クリックログ内で前記現行クエリデータに対応する被クリックカテゴリに対応するクリック特徴を含む、方法。

適用例 8：適用例 7 に記載の方法であって、前記クリック特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータが前記クリックログ内に存在するか否かを判定し、前記クリックログは、前記クエリデータおよび前記クリック特徴の間の対応関係を格納するよう適合されており、前記現行クエリデータが前記クリックログ内に存在する場合に、前記事前設定されたクリックログから前記現行クエリデータに対応する第 1 の所定の数の被クリックカテゴリを取得し、前記被クリックカテゴリは、前記クエリデータの検索後にユーザによってクリックされた対応する製品が属するカテゴリであり、前記現行クエリデータが前記クリックログ内に存在しない場合に、前記現行クエリデータに最も近い被クリッククエリデータについて前記クリックログに照会し、前記被クリッククエリデータに対応する前記第 1 の所定の数のカテゴリを前記クリック特徴として取得し、前記被クリッククエリデータは、前記現行クエリデータから最短編集距離にある前記クリックログ内のデータであること、を備える、方法。

適用例 9：適用例 1 に記載の方法であって、前記一般属性特徴は、選択ログ内で前記現行クエリデータに対応する少なくとも 1 つの被選択カテゴリに対応する選択特徴を含む、方法。

適用例 10：適用例 9 に記載の方法であって、前記選択特徴を含む前記一般属性特徴の抽出は、前記現行クエリデータが前記選択ログ内に存在するか否かを判定し、前記選択ログは、前記クエリデータおよび前記選択特徴の間の対応関係を格納するよう適合されており、前記現行クエリデータが前記選択ログ内に存在する場合に、前記選択ログから前記現行クエリデータに対応する第 1 の所定の数の被選択カテゴリを取得し、前記被選択カテゴリは、検索エンジンのマッピングテーブル内で前記現行クエリデータに関連付けられたカテゴリを表し、前記現行クエリデータが前記選択ログ内に存在しない場合に、前記現行クエリデータに最も近い被選択クエリデータについて前記選択ログに照会し、前記被選択クエリデータに対応する前記第 1 の所定の数のカテゴリを前記現行クエリデータの選択特徴として取得し、前記被選択クエリデータは、前記現行クエリデータから最短編集距離にある前記選択ログ内のデータであること、を備える、方法。

適用例 11：適用例 1 に記載の方法であって、前記一般属性特徴は、同一セッションにおける前記現行クエリデータの前の 1 つのクエリに対応する実質的に最適なカテゴリに対応するセッション特徴を含む、方法。

適用例 12：適用例 11 に記載の方法であって、前記セッション特徴を含む前記一般属性特徴の抽出は、前記同じセッションにおける前記現行クエリデータの前の先行するクエリデータを取得し、前記先行するクエリデータに対応する最適なカテゴリ分け結果が属するカテゴリを取得し、前記最適なカテゴリ分け結果が属する前記カテゴリを前記現行クエリデータのセッション特徴として設定すること、を備える、方法。

適用例 13：適用例 1 に記載の方法であって、前記一般属性特徴に少なくとも部分的に基づく、前記現行クエリデータの前記複数の確度の決定は、前記一般属性特徴を条件付き確率モデルの所与情報として設定し、前記条件付き確率モデルを用いて、前記現行クエリが前記所与情報の条件下で前記それぞれのカテゴリに属する複数の事後確率を計算し、前記複数の事後確率を前記複数の確度として設定すること、を備える、方法。

適用例 14：適用例 1 に記載の方法であって、さらに、前記複数の確度の内、既定の第 1 の閾値を満たす選択された確度と、前記選択された確度に対応するカテゴリとを特定し、前記現行クエリデータについての前記選択カテゴリを選択するために、前記複数の確度の内の前記選択された確度と、前記選択された確度に対応するカテゴリとを前記クライアントに返すこと、を備える、方法。

10

20

30

40

50

適用例 15：適用例 1 に記載の方法であって、前記選択カテゴリはユーザによって選択される、方法。

適用例 16：適用例 1 に記載の方法であって、前記選択カテゴリは、前記現行クエリデータについて計算された最大確度を有する、方法。

適用例 17：情報を検索するためのシステムであって、1 または複数のプロセッサであって、クライアントから現行クエリデータを受信し、前記現行クエリデータの一般属性特徴を抽出し、前記一般属性特徴は、複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテゴリに属する確度を示し、前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の確度を決定し、選択カテゴリ内において、前記現行クエリデータに対応する検索結果を検索し、前記選択カテゴリは、前記複数のカテゴリの 1 つであり、前記複数の確度に少なくとも部分的に基づいて選択され、前記検索結果を返すことを実行するよう構成されている 1 または複数のプロセッサと、

10

前記 1 または複数のプロセッサに接続され、前記プロセッサに命令を提供するよう構成されている 1 または複数のメモリと、を備える、システム。

適用例 18：個人の特徴を推測するためのコンピュータプログラム製品であって、前記コンピュータプログラム製品は、コンピュータ読み取り可能な記憶媒体内に組み込まれ、クライアントから現行クエリデータを受信するためのコンピュータ命令と、前記現行クエリデータの一般属性特徴を抽出するためのコンピュータ命令と、前記一般属性特徴は、複数のカテゴリに対応する前記現行クエリデータの複数の確度を計算するために用いられ、前記確度の各々は、前記現行クエリデータが前記複数のカテゴリの内のそれぞれのカテゴリに属する確度を示し、前記一般属性特徴に少なくとも部分的に基づいて、前記現行クエリデータの前記複数の確度を決定するためのコンピュータ命令と、選択カテゴリ内で、前記現行クエリデータに対応する検索結果を検索するためのコンピュータ命令と、前記選択カテゴリは、前記複数のカテゴリの 1 つであり、前記複数の確度に少なくとも部分的に基づいて選択され、前記検索結果を返すためのコンピュータ命令と、を備える、コンピュータプログラム製品。

20

【図 1】

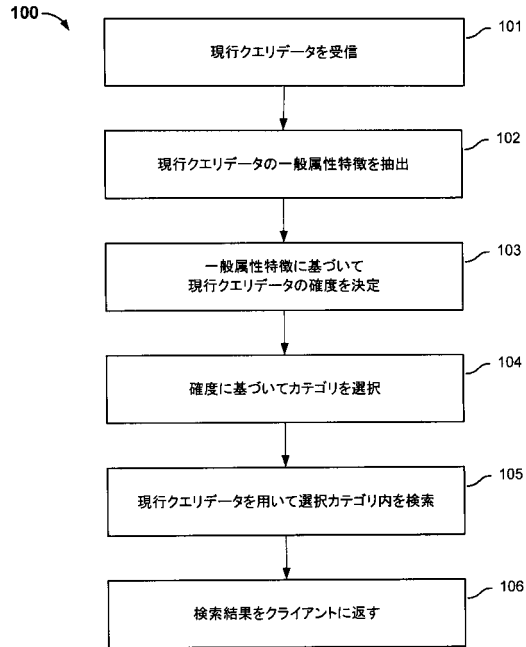


FIG. 1

【図 2】



FIG. 2

【図 3】

カテゴリ:	MP3プレーヤ	MP4プレーヤ	携帯電話
	イヤホンおよびヘッドホン	USBフラッシュドライブ	販促品

FIG. 3

【図 4】

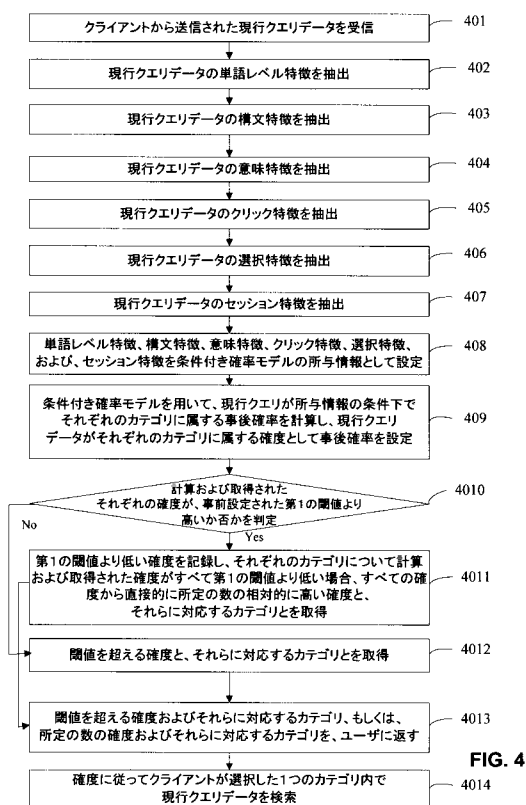


FIG. 4

【図 5】

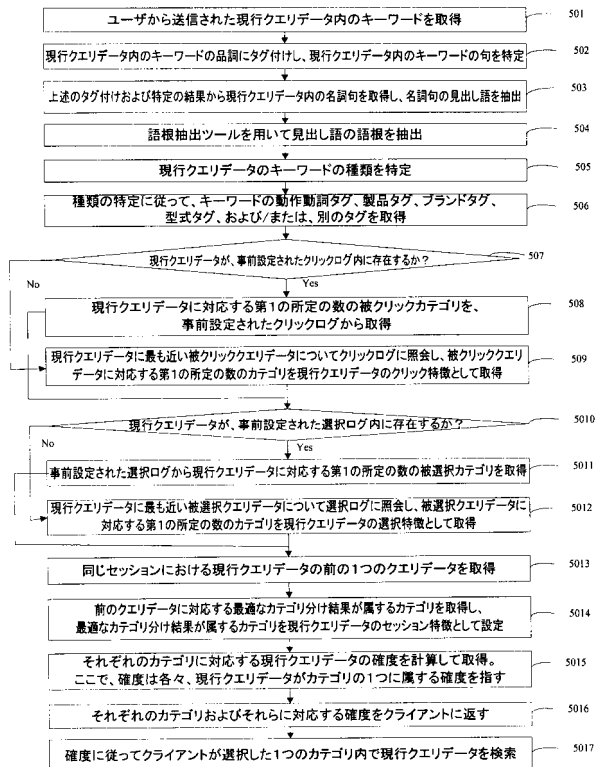


FIG. 5

【図 6】

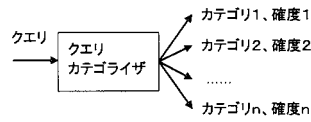


FIG. 6

【図 7】

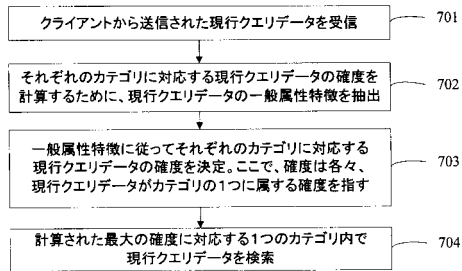


FIG. 7

【図 8】

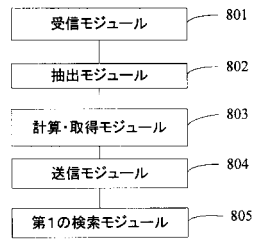


FIG. 8

【図 10】

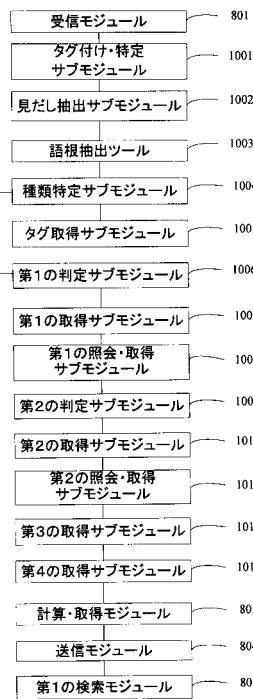


FIG. 10

【図 9】

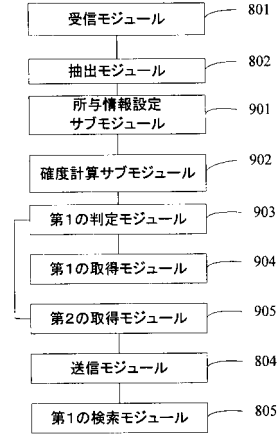


FIG. 9

【図 11】

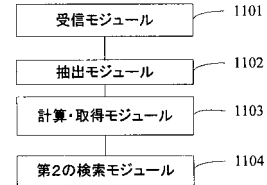


FIG. 11

フロントページの続き

- (72)発明者 カン・チャンホイ
中華人民共和国 ハンチョウ, ワーナー・ロード, ウエスト・レイク・インターナショナル・プラ
ザ, 10階, ナンバー391
- (72)発明者 ファン・ペン
中華人民共和国 ハンチョウ, ワーナー・ロード, ウエスト・レイク・インターナショナル・プラ
ザ, 10階, ナンバー391

審査官 久々宇 篤志

- (56)参考文献 特開2005-149033(JP, A)
特開2008-203933(JP, A)
特表2008-523469(JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F 17/30