

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成25年3月14日(2013.3.14)

【公表番号】特表2009-528619(P2009-528619A)

【公表日】平成21年8月6日(2009.8.6)

【年通号数】公開・登録公報2009-031

【出願番号】特願2008-557273(P2008-557273)

【国際特許分類】

G 06 F 17/30 (2006.01)

【F I】

G 06 F 17/30 3 4 0 B

【誤訳訂正書】

【提出日】平成25年1月25日(2013.1.25)

【誤訳訂正1】

【訂正対象書類名】特許請求の範囲

【訂正対象項目名】全文

【訂正方法】変更

【訂正の内容】

【特許請求の範囲】

【請求項1】

複数のユーザの直接観察された振舞いの特性及び導き出された振舞いの特性に基づいて、問い合わせに対する検索結果の、ユーザのユーザ選択を予測するユーザ振舞いコンポーネントであって、前記直接観察された振舞いの特性は、前記複数のユーザのウェブページ対話を定量化する値から測定され、前記導き出された振舞いの特性は、前記直接観察された振舞いの特性の問い合わせ独立の分布から予測された期待値からの、前記直接観察された振舞いの特性の値の偏差を表し、さらに、各問い合わせURLペア、前記問い合わせの組み合わせである各問い合わせURLペア、及び前記問い合わせから取得される前記検索結果中の対応するウェブページのURL(uniform resource locator)に応じて、前記複数のユーザおよび検索セッションにまたがる前記直接観察された振舞いの特性についての値を平均化することで、前記複数のユーザのユーザ振舞いにおけるバリエーションを軽減するユーザ振舞いコンポーネントと、

前記ユーザおよび前記複数のユーザのウェブブラウジングの対話をキャプチャ及び定量化する表示特性、ブラウジング(走査検索)特性、またはクリックスルーパー特性の少なくとも一つを備え、前記直接観察された振舞いの特性及び前記導き出された振舞いの特性を含む特性セットと、

検索結果の関連性及び順位の決定のための前記ユーザ選択を組み込む検索エンジンと
であるコンピュータ実行可能なコンポーネントを備えることを特徴とするコンピュータ実装のシステム。

【請求項2】

前記ユーザ振舞いコンポーネントは、バックグラウンドコンポーネント及び関連性コンポーネントをさらに備えることを特徴とする請求項1に記載のコンピュータ実装のシステム。

【請求項3】

機械学習コンポーネントをさらに備えることを特徴とする請求項1に記載のコンピュータ実装のシステム。

【請求項4】

前記ユーザ振舞いコンポーネントは、ユーザの振舞いのデータドリブンモデルをさらに備えることを特徴とする請求項1に記載のコンピュータ実装のシステム。

【請求項 5】

前の検索データを含むデータログをさらに備えることを特徴とする請求項4に記載のコンピュータ実装のシステム。

【請求項 6】

前記検索エンジンは、検索結果を順位付ける順位付けコンポーネントをさらに備えることを特徴とする請求項1に記載のコンピュータ実装のシステム。

【請求項 7】

前記データドリブンモデルを訓練する機械学習コンポーネントをさらに備えることを特徴とする請求項4に記載のコンピュータ実装のシステム。

【請求項 8】

検索エンジンとの対話の間、前記検索エンジンと対話を行うユーザ振舞いコンポーネントが、複数のユーザの、ユーザの振舞いを取得するステップであって、前記検索エンジンにユーザの問い合わせを送信する、ステップと、

前記問い合わせによって取得される検索結果の、ユーザのユーザ選択を予測するために、前記ユーザ振舞いコンポーネントが、前記ユーザの振舞いの直接観察された特性及び導き出された振舞いの特性の分析についての前記ユーザの振舞いを統合するステップであって、前記直接観察された特性は、前記複数のユーザの対話を定量化する値から測定され、前記導き出された特性は、前記複数のユーザの前記直接観察された特性の問い合わせ独立の分布から予測された期待値からの、前記直接観察された特性の値の偏差を表す、ステップと、

前記ユーザ振舞いコンポーネントが、各問い合わせURLペア、前記問い合わせの組み合わせである各問い合わせURLペア、及び前記問い合わせから取得される前記検索結果中の対応するウェブページのURL (uniform resource locator)に応じて、前記複数のユーザおよび検索セッションにまたがる、前記対話の前記直接観察された特性の値を平均化することで、前記複数のユーザのユーザ振舞いに関連するノイズであって、悪意のある又は不合理なブラウジング動作に対応するノイズを軽減するステップと、

前記ユーザ振舞いコンポーネントが、前記ユーザの振舞いを統合するステップ、及び前記ノイズを軽減するステップに基づいて、取り出された結果に対するユーザ選択を予測するステップと

前記検索エンジンが、前記ユーザ選択に基づいて、前記取り出された結果の順位付けを決定するステップと、

であるコンピュータ実行可能な動作を備えたことを特徴とするコンピュータ実装の方法。

【請求項 9】

前記ユーザ振舞いコンポーネントが、前記取り出された結果の順位付けのためのモデルを訓練するステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 10】

前記ユーザ振舞いコンポーネントが、前記ユーザの振舞いからモデルを自動的に生成するステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 11】

前記ユーザ振舞いコンポーネントが、前記取り出された結果とのユーザの対話に関する特性のセットを作り出すステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 12】

前記検索エンジンが、前記ユーザの振舞いを組み込むために機械学習を採用するステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 13】

前記ユーザ振舞いコンポーネントが、前記ユーザの振舞いを予測するステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 14】

前記検索エンジンが、前記取り出された結果の順位付けのために、統合されたユーザの振舞いをマイニングするステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 15】

前記ユーザ振舞いコンポーネントが、前記ユーザ選択を予測するために、前記取り出された結果との前記ユーザの対話から、直接観察された特性を採用するステップをさらに備えたことを特徴とする請求項8に記載のコンピュータ実装の方法。

【請求項 16】

実行されると、一つまたは複数のプロセッサに、

複数のユーザによるウェブブラウジングの振舞いについてのデータを取得するステップと、

前記データからユーザ選択を予測するモデルを形成するステップであって、前記モデルは直接観察された特性及び導き出された特性を含む特性セットを使用し、前記直接観察された特性は観察されたユーザのブラウジングの振舞いを定量化する値によって測定され、前記導き出された特性は、前記直接観察された特性の問い合わせ独立の分布から予測された期待値からの、前記直接観察された特性の値の偏差を表し、前記特性セットは、前記複数のユーザのウェブブラウジングの対話をキャプチャ及び定量化する表示特性、ブラウジング（走査検索）特性、またはクリックスルー特性の少なくとも一つを備え、さらに、ユーザの振舞いを予測することは、ウェブ検索問い合わせ、前記ウェブ検索問い合わせの組み合わせである各問い合わせURLペア、及びウェブ検索問い合わせから取得される前記検索結果中の対応するウェブページのURL(uniform resource locator)から取得される検索結果に含まれる複数の問い合わせURLペアのそれぞれに対して、前記複数のユーザおよび検索セッションにまたがる前記直接観察された特性についての値を平均化することを含む、ステップと、

前記ユーザ選択に基づいて前記ウェブ検索問い合わせの前記検索結果に順位を割り当てるために、前記モデルと対話するステップと、

を含む動作を実行させる、コンピュータが実行可能な命令を格納するコンピュータ可読メモリ。

【請求項 17】

実行されると、一つまたは複数のプロセッサに、

直接観察された特性及び導き出された特性の両方を使用した検索問い合わせ結果ペアに応じたユーザの振舞いをモデル化するステップ

をさらに含む動作を実行させる、コンピュータが実行可能な命令を格納する請求項16に記載のコンピュータ可読メモリ。

【請求項 18】

実行されると、一つまたは複数のプロセッサに、

問い合わせがウェブドキュメントのテキストにどのくらい近く一致するかを含む問い合わせ独立の特性、及び、ウェブドキュメントのページランク（PageRank）を含む問い合わせ独立の特性、の両方に基づいて前記検索結果を順位付けするステップ

をさらに含む動作を実行させる、コンピュータが実行可能な命令を格納する請求項16に記載のコンピュータ可読メモリ。

【誤訳訂正2】

【訂正対象書類名】明細書

【訂正対象項目名】0012

【訂正方法】変更

【訂正の内容】

【0012】

さらに、主題の工夫点のユーザの振舞いコンポーネントは、ユーザの振舞いを説明するための2つの特性タイプを有し、直接(direct)及び偏差(deviational)という名であり、

前者は直接計測された値であり、後者は対応する直接観察された特性についての（問い合わせ独立の）分布全体から予測された期待値からの差である。したがって、問い合わせ q 及び結果 r に対する特性 f の観察された値 o は、2つのコンポーネントの混合として表現でき、

$$o(q, r, f) = C(r, f) + r\text{el}(q, r, f)$$

であり、 $C(r, f)$ は r に対応する全ての問い合わせにまたがって統合された f の値についての前の「バックグラウンド」分布であり、 $r\text{el}(q, r, f)$ は、問い合わせへの結果の関連性により影響された振舞いの「関係性のある」コンポーネントである。例えば、ユーザの振舞いの関連性の予測は、所与の位置で観察されたクリックスルーフ周波からのバックグラウンド分布の減算を介してクリックスルーフ特性で取得できる。振舞いにおける個々のユーザのバリエーションの影響を軽減するために、主題の工夫点により、各問い合わせ結果ペアについて、全てのユーザおよび検索セッション(session)にまたがる特性値を平均化できる。このような集合は、さらなる構造安定性を供給でき、「うるさい(noisy)」ユーザの対話に頼らない。

【誤訳訂正3】

【訂正対象書類名】明細書

【訂正対象項目名】0024

【訂正方法】変更

【訂正の内容】

【0024】

図2は、主題の工夫点のさらなる態様を図示し、検索エンジン202は主題の工夫点の態様に従って、訓練モデル204をさらに備える。訓練モデル204は、観察された振舞い特性201及び導き出された振舞い特性203という名で、ユーザの振舞いを説明するための追加のモデルタイプをさらに備えることができる。観察された振舞い特性201は直接測定された値であり、導き出された振舞い特性203は、対応する直接観察された特性に対する全体（問い合わせ独立の）分布から予測された期待された値から導き出される。従って、問い合わせ q 及び結果 r に対する特性 f の観察された値は2つのコンポーネントの混合として表現できる：

$$o(q, r, f) = C(r, f) + r\text{el}(q, r, f)$$

ここで $C(r, f)$ は、 r に対応する全ての問い合わせに渡って統合された f の値の前の「バックグラウンド」分布であり、 $r\text{el}(q, r, f)$ は結果の関連性により影響を受けた振舞いのコンポーネントである。例えばユーザの振舞いの関連性の予測は、所与の位置で観察されたクリックスルーフ周波からバックグラウンド分布（例えばノイズ）の減算を介して、クリックスルーフ特性で取得できる。振舞いにおける個々のユーザのバリエーションの影響を軽減するために、主題の工夫点により、各問い合わせURLペアについて、全てのユーザおよび検索セッション(session)にまたがる直接の特性値を平均化できる。このようなさらなるロバスト性を供給でき、個々の「うるさい」ユーザ対話によらない。したがって問い合わせURLペアに対するユーザの振舞いは、直接観察された特性及び導き出された「修正された」特性値の両方を含む特性ベクトルにより表現できる。