

(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2012-168844  
(P2012-168844A)

(43) 公開日 平成24年9月6日(2012.9.6)

|                             |                 |             |
|-----------------------------|-----------------|-------------|
| (51) Int.Cl.                | F I             | テーマコード (参考) |
| <b>G06F 17/30 (2006.01)</b> | G06F 17/30 340Z | 5B075       |
|                             | G06F 17/30 340A |             |
|                             | G06F 17/30 330Z |             |

審査請求 有 請求項の数 5 O L (全 13 頁)

(21) 出願番号 特願2011-30619 (P2011-30619)  
(22) 出願日 平成23年2月16日 (2011.2.16)

(71) 出願人 500257300  
ヤフー株式会社  
東京都港区赤坂9丁目7番1号  
(72) 発明者 工藤 和也  
東京都港区赤坂9丁目7番1号 ヤフー株式会社内  
Fターム(参考) 5B075 PR04 QM08

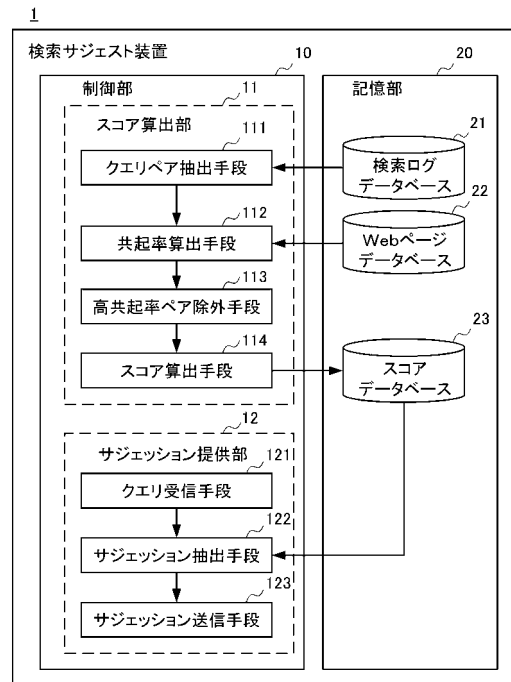
(54) 【発明の名称】 検索サジェスト装置及び方法

(57) 【要約】

【課題】 検索結果の同一性を回避しつつ次々と新たな話題を展開する娯楽性の高い検索結果を提供可能な検索クエリを自動的に提案する検索サジェスト装置及び方法を提供すること。

【解決手段】 検索サジェスト装置(1)は、検索ログデータベース(21)から抽出した検索クエリ及び再検索クエリの組み合わせの各々について検索クエリ及び再検索クエリのWebページ中の共起率を算出する共起率算出手段(112)と、算出した共起率が所定以上である組み合わせを除外する高共起率ペア除外手段(113)と、除外された組み合わせ以外の組み合わせについて検索回数が多く共起率が低いほど高いスコアを算出するスコア算出手段(114)と、ユーザ端末から検索クエリを受け付けると、当該検索クエリに対するスコアの高い再検索クエリから優先してサジェッションクエリ抽出し、ユーザ端末に送信するサジェッション送信手段(123)と、を備える。

【選択図】 図1



## 【特許請求の範囲】

## 【請求項 1】

検索クエリに対するサジェッションクエリを決定するためのスコアを予め算出するスコア算出部と、ユーザ端末から検索クエリを受け付けたことを契機として、前記ユーザ端末に前記スコア算出部が算出したスコアに基づいてサジェッションクエリを提供するサジェッション提供部と、を備えた検索サジェスト装置であって、

検索クエリと当該検索クエリに基づく検索の後に引き続き受け付けた再検索クエリとを関連付けて記憶する検索ログデータベースを更に備え、

前記スコア算出部は、

前記検索ログデータベースから、前記検索クエリと前記再検索クエリとの組み合わせを抽出し、複数の Web ページを記憶する Web ページデータベースを参照して、抽出した組み合わせの各々について当該組み合わせに係る前記検索クエリと前記再検索クエリとの前記 Web ページにおける共起率を算出する共起率算出手段と、

前記共起率算出手段が算出した前記共起率が所定以上である場合に、当該組み合わせを除外する高共起率ペア除外手段と、

前記高共起率ペア除外手段により除外された組み合わせを除いた前記検索クエリと前記再検索クエリとの組み合わせについて、当該組み合わせの検索回数が多く前記共起率が低いほど高いスコアを算出するスコア算出手段と、

を備え、

前記サジェッション提供部は、

前記ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出するサジェッション抽出手段と

前記サジェッション抽出手段が抽出した前記サジェッションクエリを前記ユーザ端末に送信するサジェッション送信手段と、

を備える検索サジェスト装置。

## 【請求項 2】

前記サジェッション抽出手段は、

前記ユーザ端末から第 1 検索クエリを受け付けたことを契機として、当該第 1 検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出し、

前記ユーザ端末から前記第 1 検索クエリに続いて第 2 検索クエリを受け付けたことを契機として、当該第 2 検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち、前記第 1 検索クエリと前記再検索クエリとに基づく前記スコア及び前記第 2 検索クエリと前記再検索クエリとに基づく前記スコアにより算出される合計スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出する、

請求項 1 に記載の検索サジェスト装置。

## 【請求項 3】

前記サジェッション抽出手段は、前記ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを一方に含む前記組み合わせのうち前記スコアの高い組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出する、

請求項 1 に記載の検索サジェスト装置。

## 【請求項 4】

前記サジェッション抽出手段は、

前記ユーザ端末から第 1 検索クエリを受け付けたことを契機として、当該第 1 検索クエリを一方に含む前記組み合わせのうち前記スコアの高い組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出し、

前記ユーザ端末から前記第 1 検索クエリに続いて第 2 検索クエリを受け付けたことを契機として、当該第 2 検索クエリを一方に含む前記組み合わせのうち、前記第 1 検索クエ

10

20

30

40

50

りを一方に含む前記組み合わせの前記スコア及び前記第2検索クエリを一方に含む前記スコアにより算出される合計スコアの高い前記組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出する、

請求項3に記載の検索サジェスト装置。

【請求項5】

検索クエリと当該検索クエリに基づく検索の後に引き続き受け付けた再検索クエリとを関連付けて記憶する検索ログデータベースを備える検索サジェスト装置が実行する、

前記検索ログデータベースから、前記検索クエリと前記再検索クエリとの組み合わせを抽出し、複数のWebページを記憶するWebページデータベースを参照して、抽出した組み合わせの各々について当該組み合わせに係る前記検索クエリと前記再検索クエリとの前記Webページにおける共起率を算出するステップと、

算出した前記共起率が所定以上である場合に、当該組み合わせを除外するステップと、

除外された組み合わせを除いた前記検索クエリと前記再検索クエリとの組み合わせについて、当該組み合わせの検索回数が多く前記共起率が低いほど高いスコアを算出するステップと、

ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先してサジェッションクエリとして抽出し、前記ユーザ端末に送信するステップと、

を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、検索サジェスト装置及び方法に関し、特に、ユーザ端末から検索クエリを受け付けると、当該検索クエリに関連する検索ワードを提案（サジェッション）する検索サジェスト装置及び方法に関する。

【背景技術】

【0002】

ネットワークを介したWWW（World Wide Web）などのドキュメントシステムにおいては、非常に多くの数のウェブページが公開されている。ユーザは、検索サイトにおいて適宜の検索クエリを入力し、検索結果として表示される複数の候補から検索条件に適合するウェブページのURL（Uniform Resource Locator）を選択し、その選択したURLへのリンクを辿ることで、所望のウェブページを閲覧できる。

【0003】

ウェブページの検索システムは、日々発展しており、近年では、入力された検索クエリに関連する情報を自動的に提案（サジェッション）することが一般的に行われている。入力された検索クエリに対して自動的に提案する情報は、検索の利便性に直接的に影響を与えるものであり、同一の意味を持つ同意語や、相互に何らかの関係性を有する関連語をインターネットの検索時に用いる検索クエリを利用して特定する方法が知られている。これらの方法によって、特定された単語の組み合わせは、同意語や関連語などではないこともあり、そのような組み合わせをいかに取り除くかに技術上の主眼がある。

例えば、非特許文献1には、検索サイトのクエリログ（検索ログ）を利用し、共通の上位語を持つ語を抽出する方法が開示されている。

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】山口雅史、大島裕明、小山聡、田中克己、著「サーチエンジンのクエリログを利用した同位語の発見」、[online]、[平成22年12月27日検索]、  
<http://www.dbsj.org/Japanese/DBSJLetter

10

20

30

40

50

s / v o l 1 5 / n o 2 / y a m a g u c h i . p d f >

【発明の概要】

【発明が解決しようとする課題】

【0005】

ところで、従来の方法により抽出される関連語は、例えば、検索クエリ「高校野球」に対して関連語「甲子園」などのように一見して関連することが明確な単語であるため、検索結果が同一になり易く、ユーザに次々と新たな話題を展開する娯楽性の高い検索結果を提供することができなかった。

【0006】

そこで、本発明は、検索結果の同一性を回避しつつ次々と新たな話題を展開する娯楽性の高い検索結果を提供可能な検索クエリを自動的に提案する検索サジェスト装置及び方法を提供することを目的とする。

【課題を解決するための手段】

【0007】

検索システムを利用するユーザは、所望する情報を検索するために一連の検索操作において、様々な検索クエリを入力することがある。このように入力される様々な検索クエリのそれぞれは、一見して関連することが明確な単語もあれば、一見しただけでは関連性がない単語もある。これらの一見すると関連がない単語であっても、ユーザは、一定の検索意図で検索を行っており、一定の検索意図での関連性を有していることを本発明者は見出し、本発明を完成するに至った。

【0008】

(1) 検索クエリに対するサジェッションクエリを決定するためのスコアを予め算出するスコア算出部と、ユーザ端末から検索クエリを受け付けたことを契機として、前記ユーザ端末に前記スコア算出部が算出したスコアに基づいてサジェッションクエリを提供するサジェッション提供部と、を備えた検索サジェスト装置であって、検索クエリと当該検索クエリに基づく検索の後に引き続き受け付けた再検索クエリとを関連付けて記憶する検索ログデータベースを更に備え、前記スコア算出部は、前記検索ログデータベースから、前記検索クエリと前記再検索クエリとの組み合わせを抽出し、複数のWebページを記憶するWebページデータベースを参照して、抽出した組み合わせの各々について当該組み合わせに係る前記検索クエリと前記再検索クエリとの前記Webページにおける共起率を算出する共起率算出手段と、前記共起率算出手段が算出した前記共起率が所定以上である場合に、当該組み合わせを除外する高共起率ペア除外手段と、前記高共起率ペア除外手段により除外された組み合わせを除いた前記検索クエリと前記再検索クエリとの組み合わせについて、当該組み合わせの検索回数が多く前記共起率が低いほど高いスコアを算出するスコア算出手段と、を備え、前記サジェッション提供部は、前記ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出するサジェッション抽出手段と、前記サジェッション抽出手段が抽出した前記サジェッションクエリを前記ユーザ端末に送信するサジェッション送信手段と、を備える検索サジェスト装置。

【0009】

(1)の検索サジェスト装置によれば、スコア算出部が予め検索クエリと再検索クエリとの組み合わせの検索回数からスコアを算出すると、サジェッション提供部は、このスコアに基づいて、ユーザ端末から受け付けた検索クエリに対するサジェッションクエリを送信する。これにより、検索システムを利用する多数のユーザの検索意図を反映したサジェッションクエリをユーザ端末に送信することができる。

このとき、スコア算出部の高共起率ペア除外手段は、検索クエリと再検索クエリとのWebページにおける共起率が所定以上の組み合わせを除外するとともに、スコア算出手段は、共起率が低いほど高いスコアを算出する。これにより、ユーザ端末から受け付けた検索クエリと共起率が低い関係にあるサジェッションクエリがユーザ端末に送信される。そ

10

20

30

40

50

の結果、サジェッションクエリによる検索と検索クエリによる検索とで、同一の検索結果が得られることを回避することができ、ユーザ端末に次々と新たな話題を展開する娯楽性の高い検索結果を得るためのサジェッションクエリを自動的に提案することができる。

【0010】

(2) 前記サジェッション抽出手段は、前記ユーザ端末から第1検索クエリを受け付けたことを契機として、当該第1検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出し、前記ユーザ端末から前記第1検索クエリに続いて第2検索クエリを受け付けたことを契機として、当該第2検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち、前記第1検索クエリと前記再検索クエリとに基づく前記スコア及び前記第2検索クエリと前記再検索クエリとに基づく前記スコアにより算出される合計スコアの高い再検索クエリから優先して前記サジェッションクエリとして抽出する、(1)に記載の検索サジェスト装置。

10

【0011】

(2)の検索サジェスト装置によれば、ユーザ端末から第1検索クエリに続いて第2検索クエリを受け付けた場合には、サジェッションクエリを抽出するためのスコアに第1検索クエリについて算出されるスコアを加味することができる。その結果、同一の(第2)検索クエリに対して、異なるサジェッションクエリを提案することができ、ユーザ端末に次々と新たな話題を展開する娯楽性の高い検索結果を提供することができる。

【0012】

(3) 前記サジェッション抽出手段は、前記ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを一方に含む前記組み合わせのうち前記スコアの高い組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出する、(1)に記載の検索サジェスト装置。

20

【0013】

(1)の検索サジェスト装置では、検索クエリと再検索クエリとの順序を考慮した上でサジェッションクエリを提案することとしていたが、(3)の検索サジェスト装置によれば、検索クエリと再検索クエリとの順序を考慮することなくサジェッションクエリを提案することができる。例えば、「検索クエリ(単語A) - 再検索クエリ(単語B) : スコア(300)」、「検索クエリ(単語A) - 再検索クエリ(単語C) : スコア(200)」、「検索クエリ(単語C) - 再検索クエリ(単語A) : スコア(200)」であり、ユーザ端末から検索クエリ(単語A)を受け付けた場合、(1)の検索サジェスト装置によれば、単語Bを優先してサジェッションクエリとして提案する一方で、(3)の検索サジェスト装置によれば、単語Cを優先してサジェッションクエリとして提案することができる。このような(3)の検索サジェスト装置によっても、検索結果の同一性を回避しつつ次々と新たな話題を展開する娯楽性の高い検索結果を得るための検索クエリを自動的に提案することができる。

30

【0014】

(4) 前記サジェッション抽出手段は、前記ユーザ端末から第1検索クエリを受け付けたことを契機として、当該第1検索クエリを一方に含む前記組み合わせのうち前記スコアの高い組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出し、前記ユーザ端末から前記第1検索クエリに続いて第2検索クエリを受け付けたことを契機として、当該第2検索クエリを一方に含む前記組み合わせのうち、前記第1検索クエリを一方に含む前記組み合わせの前記スコア及び前記第2検索クエリを一方に含む前記スコアにより算出される合計スコアの高い前記組み合わせに含まれるクエリから優先して前記サジェッションクエリとして抽出する、(3)に記載の検索サジェスト装置。

40

【0015】

(4)の検索サジェスト装置によれば、(2)の検索サジェスト装置と同様に、同一の(第2)検索クエリに対して、異なるサジェッションクエリを提案することができ、ユーザ端末に次々と新たな話題を展開する娯楽性の高い検索結果を提供することができる。こ

50

のとき、(4)の検索サジェスト装置によれば、(2)の検索サジェスト装置とは異なり、検索クエリと再検索クエリとの順序を考慮することがない。

【0016】

(5) 検索クエリと当該検索クエリに基づく検索の後に引き続き受け付けた再検索クエリとを関連付けて記憶する検索ログデータベースを備える検索サジェスト装置が実行する、前記検索ログデータベースから、前記検索クエリと前記再検索クエリとの組み合わせを抽出し、複数のWebページを記憶するWebページデータベースを参照して、抽出した組み合わせの各々について当該組み合わせに係る前記検索クエリと前記再検索クエリとの前記Webページにおける共起率を算出するステップと、算出した前記共起率が所定以上である場合に、当該組み合わせを除外するステップと、除外された組み合わせを除いた前記検索クエリと前記再検索クエリとの組み合わせについて、当該組み合わせの検索回数が高く前記共起率が低いほど高いスコアを算出するステップと、ユーザ端末から検索クエリを受け付けたことを契機として、当該検索クエリを検索クエリとして含む前記組み合わせに含まれる再検索クエリのうち前記スコアの高い再検索クエリから優先してサジェッションクエリとして抽出し、前記ユーザ端末に送信するステップと、を含む方法。

10

【0017】

(5)の方法によれば、(1)の検索サジェスト装置と同様の効果を奏する。

【発明の効果】

【0018】

本発明によれば、検索結果の同一性を回避しつつ次々と新たな話題を展開する娯楽性の高い検索結果を得るための検索クエリを自動的に提案することができる。

20

【図面の簡単な説明】

【0019】

【図1】本発明の検索サジェスト装置の機能構成を示すブロック図である。

【図2】検索クエリと再検索クエリとの関係を示す図である。

【図3】検索ログデータベースを示す図である。

【図4】Webページデータベースを示す図である。

【図5】スコアデータベースを示す図である。

【図6】本発明のスコア算出処理を示すフローチャートである。

【図7】本発明のサジェッション処理を示すフローチャートである。

30

【図8】サジェッションクエリ提供画面を示す図である。

【図9】別実施形態のサジェッションクエリ提供画面を示す図である。

【発明を実施するための形態】

【0020】

以下、本発明の実施形態について図面を参照して説明する。

【0021】

[検索サジェスト装置1の構成]

図1は、検索サジェスト装置1の機能構成を示すブロック図である。検索サジェスト装置1は、制御部10と記憶部20とを含んで構成され、ネットワークを介してユーザ端末と通信可能に接続される。

40

【0022】

制御部10は、ユーザ端末から受け付けた検索クエリのログ(検索ログ)から、各検索クエリ間の関連度を示すスコアを算出するスコア算出部11と、ユーザ端末から受け付けた検索クエリに対して、スコア算出部11により算出されたスコアの高い検索クエリを提案(サジェッション)するサジェッション提供部12と、を含む。

初めに、スコア算出部11及び記憶部20について説明する。スコア算出部11は、クエリペア抽出手段111と、共起率算出手段112と、高共起率ペア除外手段113と、スコア算出手段114と、を含む。また、記憶部20は、検索ログデータベース21と、Webページデータベース22と、スコアデータベース23と、を含む。

【0023】

50

検索ログデータベース21は、ネットワークを介して複数のユーザ端末から受信した検索ログを記憶する。検索ログとは、ユーザ端末による一連の検索操作の記録（ログ）であり、ユーザ端末から受け付けた検索クエリ、当該検索クエリによる検索回数及び当該検索クエリによる検索でヒットしたWebページ数（検索件数）に加えて、検索クエリに基づく検索の後に引き続き受け付けた再検索クエリ及び検索クエリの後に行われた再検索クエリによる検索回数などを含む。ここで、検索クエリと再検索クエリとの関係について、図2を参照して具体的に説明する。

#### 【0024】

図2(1)に示す検索画面1において、検索入力欄201に「単語A」が入力され、「単語A」を検索クエリとする検索が行われると、図2(2)に示す検索画面2が表示される。図2(2)では、検索入力欄202に加え、検索結果欄211が表示され、検索結果欄211には、「単語A」の検索結果が表示される。このとき、検索結果欄211の結果内容に満足しないユーザは、引き続き検索を行うことがある。図2(2)では、検索入力欄202に（自動的に挿入されている「単語A」を削除した後）「単語B」が入力され、「単語B」を検索クエリとする検索が引き続き行われている。

10

このような検索クエリ「単語A」と、検索クエリ「単語A」に基づく検索結果が表示された画面において再度入力された検索クエリ「単語B」との関係が、検索クエリ「単語A」及び再検索クエリ「単語B」の代表的な一例である。より具体的には、例えば、検索画面1において、検索入力欄201に「サッカー」と入力され、「サッカー」を検索クエリとする検索が行われ、この検索結果画面において、検索入力欄201に入力されている「サッカー」が削除された後に「ワールドカップ」が入力され、「ワールドカップ」を検索クエリとする検索が行われる場合を考えると、検索クエリは「サッカー」となり、再検索クエリは「ワールドカップ」となる。

20

なお、「単語A」を検索クエリとする検索画面2の検索結果欄211に表示されている各Webページを参照した後に検索画面2に戻り検索クエリ「単語B」が入力された場合であっても、「単語A」と「単語B」の関係は同じように扱われる。

更に、図2(2)において「単語B」を検索クエリとする検索が行われると、「単語B」の検索結果を表示する検索結果欄212と、検索入力欄203とを含む検索画面3（図2(3)）が表示される。このとき、検索入力欄203に「単語C」が入力され検索が行われた場合には、検索クエリ「単語B」再検索クエリ「単語C」となる。

30

#### 【0025】

検索ログデータベース21は、検索クエリごとに、再検索クエリ及び検索クエリの後に行われた再検索クエリによる検索回数などを関連付けて記憶する。図3を参照して、例えば、検索クエリ「サッカー選手A」の後に再検索クエリ「サッカー大会B」が検索された数（検索クエリ「サッカー選手A」と再検索クエリ「サッカー大会B」との組み合わせの検索回数）は「2121回」であり、検索クエリ「サッカー選手A」の後に再検索クエリ「女優C」が検索された数（検索クエリ「サッカー選手A」と再検索クエリ「女優C」との組み合わせの検索回数）は「1801回」である。

#### 【0026】

図1に戻り、クエリペア抽出手段111は、検索ログデータベース21から、検索クエリと再検索クエリとの組み合わせを抽出する。例えば、クエリペア抽出手段111は、検索クエリ「サッカー選手A」と再検索クエリ「サッカー大会B」との組み合わせや、検索クエリ「サッカー選手A」と再検索クエリ「女優C」との組み合わせを検索ログデータベース21から抽出する。

40

#### 【0027】

共起率算出手段112は、クエリペア抽出手段111が抽出した検索クエリと再検索クエリとの組み合わせの各々について、Webページデータベース22を参照して検索クエリと再検索クエリとの共起率を算出する。ここで、共起率とは、ある2つの単語（検索クエリと再検索クエリ）がある範囲（Webページ）で同時に出現する頻度をいい、一例としては、以下の式により算出することができる。

50

## 【数 1】

$$X = \frac{C}{(A+B-C)}$$

X : 「検索クエリ」と「再検索クエリ」との共起率

A : 「検索クエリ」の検索件数

B : 「再検索クエリ」の検索件数

C : 「検索クエリ AND 再検索クエリ」の検索件数

なお、共起率の算出は、上記式に限られるものではなく、同一の Web ページ内における両単語間の距離などの任意の情報を用いて算出することとしてもよい。

10

## 【0028】

図 4 を参照して、Web ページデータベース 22 は、ネットワーク上の Web ページを記憶するデータベースであって、ページ ID (例えば、URL) に対応付けて Web ページの内容を記憶する。

ここで、例えば、検索クエリ「サッカー選手 A」と再検索クエリ「サッカー大会 B」とは、ページ ID「1」の Web ページ内に同時に出現しており共起率が高く算出される可能性がある。一方、検索クエリ「花火大会 X」と再検索クエリ「ゲームサイト Z」とは、同一の Web ページ内に出現することが少なく共起率が低く算出される可能性がある。

## 【0029】

高共起率ペア除外手段 113 は、共起率算出手段 112 が算出した共起率が所定以上である場合に、当該組み合わせを除外する。検索クエリと再検索クエリとの共起率が高い場合、検索クエリによる検索と再検索クエリによる検索とでは検索結果が同一になり易い。そこで、本発明の検索サジェスト装置 1 では、検索クエリと再検索クエリとの共起率が所定以上の組み合わせを除外し、同一の検索結果が得られることを回避することとしている。なお、除外する共起率の閾値は、管理者が任意に設定することができる。

20

## 【0030】

スコア算出手段 114 は、高共起率ペア除外手段 113 により除外された組み合わせを除いた検索クエリと再検索クエリとの組み合わせについて、検索クエリ・再検索クエリ間の関連度を示すスコアを算出する。このとき、スコア算出手段 114 は、検索クエリと再検索クエリとの組み合わせの検索回数が多く共起率が低いほど高いスコアを算出する。

30

## 【0031】

スコアデータベース 23 は、図 5 に示すように、スコア算出手段 114 が算出した検索クエリと再検索クエリとの関連度を示すスコアを記憶する。

ここで、図 3 を参照して、検索クエリ「花火大会 X」に対する再検索クエリ「ゲームサイト Z」の検索回数は「862回」であり、同じ検索クエリ「花火大会 X」に対する再検索クエリ「司会者 Y」の検索回数「1432回」よりも少ない。一方、図 4 を参照して、再検索クエリ「ゲームサイト Z」は、検索クエリ「花火大会 X」との同一の Web ページ内に出現することが少なく共起率が低い。再検索クエリ「司会者 Y」は、検索クエリ「花火大会 X」と同一の Web ページ内 (ページ ID「4」) に出現することがあり共起率が高い。そのため、本実施形態では、検索クエリ「花火大会 X」に対して、検索回数の少ない「ゲームサイト Z」の方が高いスコアが算出されている (図 5)。これにより、一見しただけでは関連性がない (共起率が低い) がユーザの一定の検索意図 (再検索クエリとして複数回検索されている) を反映した関連性を評価するスコアを算出することができる。

40

## 【0032】

続いて、ユーザ端末から受け付けた検索クエリに対するサジェッションクエリを提案するサジェッション提供部 12 について説明する。サジェッション提供部 12 は、クエリ受信手段 121 と、サジェッション抽出手段 122 と、サジェッション送信手段 123 と、を含む。

## 【0033】

50



クエリ受信手段 1 2 1 は、ユーザにより入力された検索クエリを、ネットワークを介してユーザ端末から受信する。

【 0 0 3 4 】

サジェッション抽出手段 1 2 2 は、スコアデータベース 2 3 を参照して、クエリ受信手段 1 2 1 が受信した検索クエリに対応する再検索クエリのうち、スコアの高い再検索クエリから順に所定数の再検索クエリをサジェッションクエリとして抽出する。例えば、ユーザ端末から検索クエリとして「花火大会 X」を受信した場合には、サジェッション抽出手段 1 2 2 は、「花火大会 X」を検索クエリとして含む組み合わせに含まれる再検索クエリのうち、スコアの高い再検索クエリから順にサジェッションクエリとして抽出する。図 5 を参照すると、最もスコアの高い「ゲームサイト Z」を最も優先して抽出し、次にスコアの

10

【 0 0 3 5 】

サジェッション送信手段 1 2 3 は、サジェッション抽出手段 1 2 2 が抽出したサジェッションクエリをネットワークを介してユーザ端末に送信する。このとき、サジェッション送信手段 1 2 3 は、サジェッション抽出手段 1 2 2 が抽出した優先順位で表示されるようにサジェッションクエリをユーザ端末に送信する。

【 0 0 3 6 】

[ 検索サジェスト装置 1 の構成 ]

以上説明した検索サジェスト装置 1 のハードウェアは、1 又は複数の一般的なコンピュータによって構成することができる。一般的なコンピュータは、例えば、制御部 1 0 として、中央処理装置 (CPU) を備える他、記憶部 2 0 として、メモリ (RAM、ROM)、ハードディスク (HDD) 及び光ディスク (CD、DVD など) を、ネットワーク通信装置として、各種有線及び無線 LAN 装置を、表示装置として、例えば、液晶ディスプレイ、プラズマディスプレイなどの各種ディスプレイを、入力装置として、例えば、キーボード及びポインティング・デバイス (マウス、トラッキングボールなど) を適宜備え、これらは、バスラインにより接続されている。このような一般的なコンピュータにおいて、CPU は、検索サジェスト装置 1 を統括的に制御し、各種プログラムを適宜読み出して実行することにより、上述したハードウェアと協働し、本発明に係る各種機能を実現している。

20

【 0 0 3 7 】

[ 検索サジェスト装置の処理 ]

続いて、図 6 及び図 7 を参照して、検索サジェスト装置 1 の処理の流れについて説明する。

30

【 0 0 3 8 】

[ スコア算出処理 ]

初めに、図 6 を参照して、スコア算出処理について説明する。図 6 は、検索サジェスト装置 1 により行われるスコア算出処理の流れを示すフローチャートである。

【 0 0 3 9 】

S 1 : クエリペア抽出手段 1 1 1 は、検索ログデータベース 2 1 から、検索クエリと再検索クエリとの組み合わせを抽出する。

40

S 2 : 続いて、共起率算出手段 1 1 2 は、S 1 で抽出した検索クエリと再検索クエリとの組み合わせの各々について、検索クエリと再検索クエリとの共起率を算出する。

S 3 : 続いて、高共起率ペア除外手段 1 1 3 は、S 2 で算出した共起率が所定以上である組み合わせを除外する。

S 4 : 続いて、スコア算出手段 1 1 4 は、S 1 で抽出した組み合わせのうち、S 3 で除外された組み合わせを除いた組み合わせについて、検索クエリ・再検索クエリ間の関連度を示すスコアを算出し、スコアデータベース 2 3 に記憶し、処理を終了する。このとき、スコア算出手段 1 1 4 は、検索クエリと再検索クエリとの組み合わせの検索回数が多いほど高いスコアを算出するとともに、共起率が低いほど高いスコアを算出する。

【 0 0 4 0 】

50

### [ サジェッション処理 ]

続いて、図7を参照して、サジェッション処理について説明する。図7は、検索サジェスト装置1により行われるサジェッション処理の流れを示すフローチャートである。

#### 【0041】

S11：クエリ受信手段121は、ユーザにより入力された検索クエリを、ネットワークを介してユーザ端末から受信する。

S12：続いて、サジェッション抽出手段122は、スコアデータベース23を参照して、S11で受信した検索クエリに対するサジェッションクエリを抽出する。このとき、サジェッション抽出手段122は、S11で受信した検索クエリに対応する再検索クエリのうち、スコアの高い再検索クエリから順に所定数の再検索クエリをサジェッションクエリとして抽出する。

10

S13：続いて、サジェッション送信手段123は、S12で抽出したサジェッションクエリをネットワークを介してユーザ端末に送信する。これにより、ユーザ端末には、一見しただけでは関連性がないがユーザの一定の検索意図を反映したサジェッションクエリが送信される。

#### 【0042】

ユーザ端末へのサジェッションクエリの提案について、図8を参照して具体的に説明する。検索入力欄204に検索クエリ「花火大会X」が入力されると、検索クエリ「花火大会X」との間のスコアの高い再検索クエリ「ゲームサイトZ」、「司会者Y」などがサジェッションクエリとして抽出され、スコアの高い順にサジェッション表示欄221に表示される。

20

ユーザは、サジェッション表示欄221に表示されたサジェッションクエリを選択することで、一見しただけでは関連性がないがユーザの一定の検索意図を反映した検索クエリを容易に選択することができる。このとき、表示されるサジェッションクエリは、検索クエリ・再検索クエリの関係にあるクエリであり、その組み合わせの検索回数が多いほど優先して表示されるため、関係性が何もないクエリを提案することを防止することができる。また、共起率が低い再検索クエリを優先してサジェッションクエリとして提案するため、サジェッションクエリに基づいて検索を行った際に同一の検索結果となってしまうことを回避することができ、話題性に富んだ娯楽性の高い検索結果を得ることができる。

30

#### 【0043】

(変形形態)

以上、本発明の実施形態について説明したが、本発明は上述した実施形態に限るものではない。また、本発明の実施形態に記載された効果は、本発明から生じる最も好適な効果を列挙したに過ぎず、本発明による効果は、本発明の実施形態に記載されたものに限定されるものではない。

#### 【0044】

上記実施形態では、サジェッションクエリの抽出に際し、ユーザ端末から受け付けた1の検索クエリのみを用いることとしているが、これに限られるものではない。例えば、ユーザ端末から1つ目の検索クエリ(第1検索クエリ)を受け付けた後、2つ目の検索クエリ(第2検索クエリ)を受け付けた場合、第2検索クエリに対するサジェッションクエリの抽出では、第1検索クエリに対するスコアを参照することとしてもよい。

40

#### 【0045】

図9を参照して、ユーザ端末から第1検索クエリ「サッカー大会B」に続いて第2検索クエリ「花火大会X」を受け付けた場合のサジェッションクエリを例にとって説明する。

#### 【0046】

図9(1)を参照して、第2検索クエリ「花火大会X」との間でスコアの高い再検索クエリは、「ゲームサイトZ(スコア:515)」、「司会者Y(スコア:265)」である。そのため、上記実施形態のようにユーザ端末から受け付けた1の検索クエリを用いてサジェッションクエリを抽出した場合には、ユーザ端末には「ゲームサイトZ」「司会者Y」の順にサジェッションクエリが提案される(図8)。

50

## 【0047】

ここで、第1検索クエリ「サッカー大会B」に対して再検索クエリ「司会者Y」にはスコア(538)が算出されていたとする(図9(1))。このとき、サジェッション抽出手段122は、この第1検索クエリ「サッカー大会B」に対するスコア(538)も加味した上で第2検索クエリ「花火大会X」に対するサジェッションクエリを抽出する。その結果、第2検索クエリ「花火大会X」に対するスコアが第1検索クエリ「サッカー大会B」に対するスコアで補正され、図9(2)に示すように、ユーザ端末には、検索入力欄205に入力された第2検索クエリ「花火大会X」に対して、「司会者Y」「ゲームサイトZ」の順にサジェッションクエリを提案するサジェッション表示欄222が表示される。

## 【0048】

なお、サジェッション抽出手段122は、第1検索クエリに対するスコアを任意に用いることができ、第1検索クエリに対するスコアをそのまま第2検索クエリに対するスコアに加算することとしてもよく、一定の重み付けをした上で(例えば、第1検索クエリに対するスコアを2分の1にした上で)第2検索クエリに対するスコアに加算することとしてもよい。また、サジェッション抽出手段122は、2つの検索クエリに基づいてサジェッションクエリを抽出するだけでなく、3つ以上の複数の検索クエリに基づいてサジェッションクエリを抽出することとしてもよい。この場合においても、それぞれの検索クエリに対するスコアをそのまま加算することとしてもよく、一定の重み付けをした上で(例えば、直近の検索クエリに対するスコアを重視した上で)スコアを加算することとしてもよい。

## 【0049】

これにより、同一の第2検索クエリに対して、第1検索クエリに応じて異なるサジェッションクエリを提案することができ、ユーザ端末に話題性に富んだ娯楽性の高い検索結果を提供することができる。

## 【0050】

また、上記実施形態では、スコア算出手段114は、検索クエリ及び再検索クエリの順序を考慮した上でスコアを算出している。すなわち、検索クエリ「花火大会X」再検索クエリ「ゲームサイトZ」に対してスコアを算出する一方で、検索クエリ「ゲームサイトZ」再検索クエリ「花火大会X」に対して別のスコアを算出しているが、これに限られるものではなく、スコア算出手段114は、検索クエリ及び再検索クエリの順序に関係なく1

のスコアを算出することとしてもよい。  
具体的には、スコア算出手段114は、検索クエリ「花火大会X」再検索クエリ「ゲームサイトZ」の検索回数、検索クエリ「ゲームサイトZ」再検索クエリ「花火大会X」の検索回数、及び「花火大会X」「ゲームサイトZ」の共起率から、「花火大会X」「ゲームサイトZ」の組み合わせについてのスコアを算出し、サジェッション抽出手段122は、ユーザ端末から検索クエリを受け付けると、当該検索クエリを一方に含む組み合わせのうちスコアの高い組み合わせに含まれるもう一方のクエリから優先してサジェッションクエリとして抽出することとしてもよい。

この場合においても同様に、サジェッション抽出手段122は、ユーザ端末から受け付けた1の検索クエリのみを用いてサジェッションクエリを抽出することとしてもよく、2つ又は複数の検索クエリを用いてサジェッションクエリを抽出することとしてもよい。

## 【0051】

また、上記実施形態で示した検索クエリと再検索クエリとの関係は一例にすぎず、図2に示した関係に限られない。再検索クエリは、検索クエリに基づく検索の後に引き続き受け付けられた検索クエリをいい、例えば、所定時間内において同じユーザ端末から受け付けた検索クエリを再検索クエリとすることとしてもよい。

## 【符号の説明】

## 【0052】

- 1 検索サジェスト装置
- 10 制御部

10

20

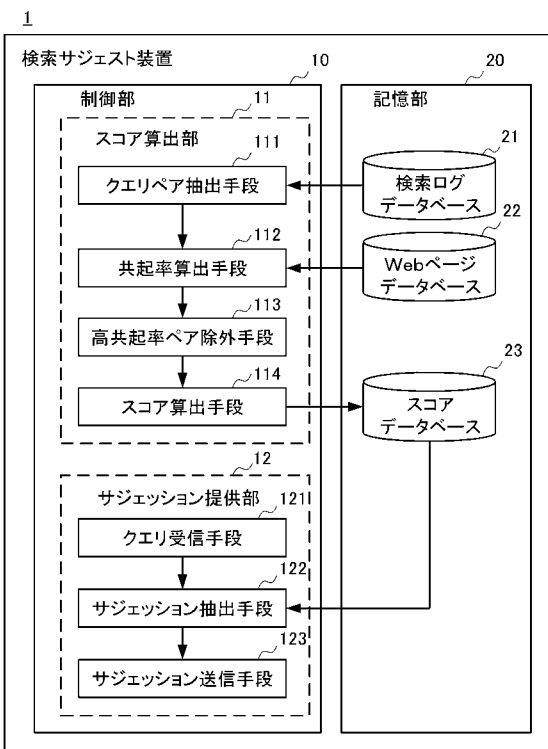
30

40

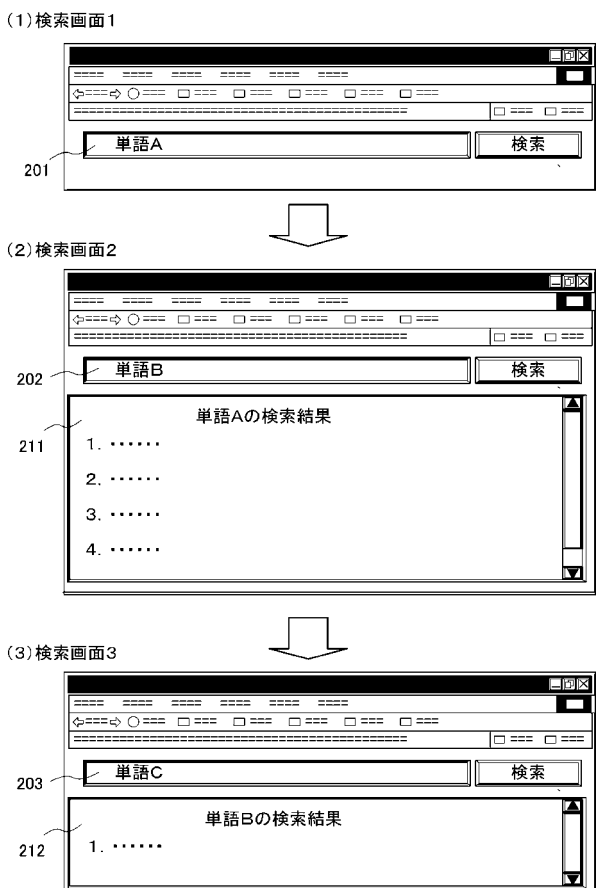
50

- 1 1 スコア算出部
- 1 1 1 クエリペア抽出手段
- 1 1 2 共起率算出手段
- 1 1 3 高共起率ペア除外手段
- 1 1 4 スコア算出手段
- 1 2 サジェッション提供部
- 1 2 1 クエリ受信手段
- 1 2 2 サジェッション抽出手段
- 1 2 3 サジェッション送信手段
- 2 0 記憶部
- 2 1 検索ログデータベース
- 2 2 Webページデータベース
- 2 3 スコアデータベース

【 図 1 】



【 図 2 】



【 図 3 】

検索ログデータベース21

| 検索クエリ   | 再検索クエリ1 |      | 再検索クエリ2 |      | ... |
|---------|---------|------|---------|------|-----|
|         | クエリ     | 検索回数 | クエリ     | 検索回数 |     |
| サッカー選手A | サッカー大会B | 2421 | 女優C     | 1801 | ... |
| 花火大会X   | 司会者Y    | 1432 | ゲームサイトZ | 862  | ... |
| ...     | ...     | ...  | ...     | ...  | ... |

【 図 4 】

Webページデータベース22

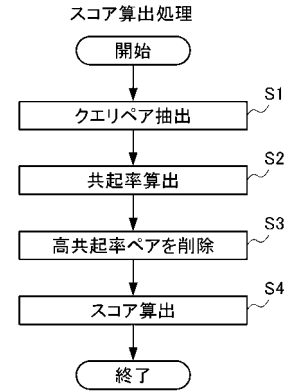
| ページID | 内容                         |
|-------|----------------------------|
| 1     | サッカー選手Aがサッカー大会Bでハットトリック... |
| 2     | サッカー選手Aが女優Cと真夜中の密会デート...   |
| 3     | 夏といえば花火！東京の花火大会Xは大盛り上がり... |
| 4     | 司会者Yが花火大会Xを盛り上げる...        |
| 5     | 暑い夏は、ゲームサイトZでゲームを楽しもう...   |
| ...   | ...                        |

【 図 5 】

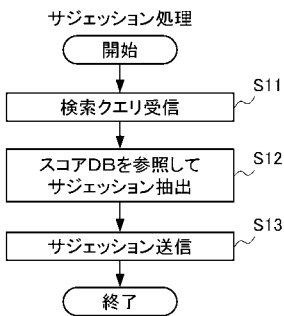
スコアデータベース23

| 検索クエリ   | 再検索クエリ  | スコア |
|---------|---------|-----|
| 花火大会X   | ゲームサイトZ | 515 |
| 花火大会X   | 司会者Y    | 265 |
| サッカー選手A | 女優C     | 288 |
| ゲームサイトZ | 花火大会X   | 422 |
| ...     | ...     | ... |

【 図 6 】

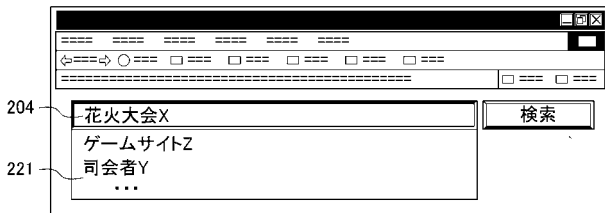


【 図 7 】



【 図 8 】

サジェッションクエリ提供画面



【 図 9 】

(1)スコアデータベース23

| 検索クエリ   | 再検索クエリ  | スコア |
|---------|---------|-----|
| 花火大会X   | ゲームサイトZ | 515 |
| 花火大会X   | 司会者Y    | 265 |
| サッカー選手A | 女優C     | 288 |
| ...     | ...     | ... |
| サッカー大会B | 司会者Y    | 538 |
| ...     | ...     | ... |

(2)サジェッションクエリ提供画面

(第2検索クエリ:花火大会X(第1検索クエリ:サッカー大会B))

