



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201403360 A

(43)公開日：中華民國 103 (2014) 年 01 月 16 日

(21)申請案號：101142220

(22)申請日：中華民國 101 (2012) 年 11 月 13 日

(51)Int. Cl. : G06F17/30 (2006.01)

G06F17/20 (2006.01)

G06Q30/00 (2012.01)

(30)優先權：2012/07/05

中國大陸

201210233248.3

(71)申請人：阿里巴巴集團服務有限公司 (香港地區) ALIBABA GROUP SERVICES LIMITED
(HK)

香港

(72)發明人：韓小梅 (CN)；宋超 (CN)；韋禕 (CN)

(74)代理人：林志剛

申請實體審查：無 申請專利範圍項數：10 項 圖式數：4 共 33 頁

(54)名稱

搜索結果的產生方法及裝置

(57)摘要

本發明提供了一種搜索結果的產生方法和裝置，所述方法包括：依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。本發明可以減少向伺服器發送請求的次數，減輕伺服器的負載。

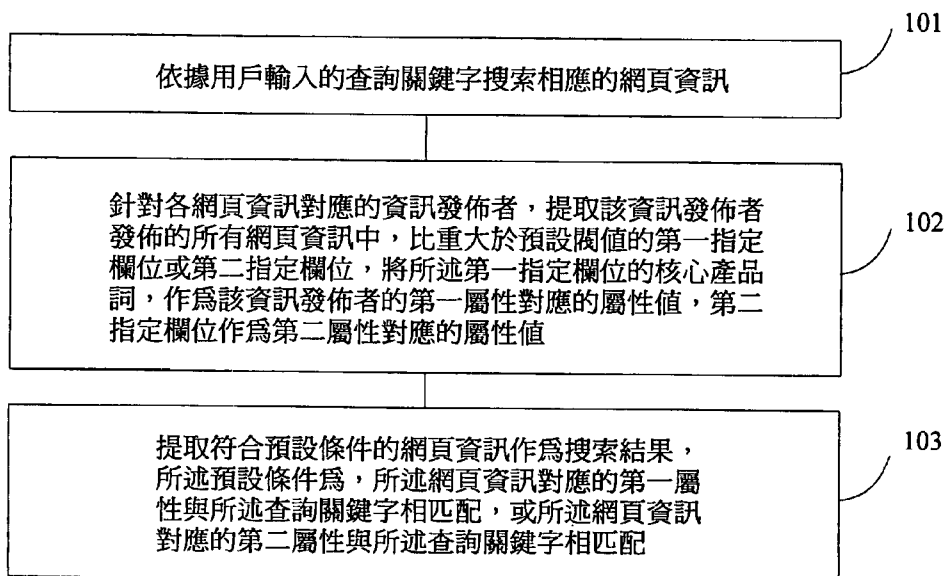


圖 1



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201403360 A

(43)公開日：中華民國 103 (2014) 年 01 月 16 日

(21)申請案號：101142220

(22)申請日：中華民國 101 (2012) 年 11 月 13 日

(51)Int. Cl. : G06F17/30 (2006.01)

G06F17/20 (2006.01)

G06Q30/00 (2012.01)

(30)優先權：2012/07/05

中國大陸

201210233248.3

(71)申請人：阿里巴巴集團服務有限公司 (香港地區) ALIBABA GROUP SERVICES LIMITED

(HK)

香港

(72)發明人：韓小梅 (CN)；宋超 (CN)；韋禕 (CN)

(74)代理人：林志剛

申請實體審查：無 申請專利範圍項數：10 項 圖式數：4 共 33 頁

(54)名稱

搜索結果的產生方法及裝置

(57)摘要

本發明提供了一種搜索結果的產生方法和裝置，所述方法包括：依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。本發明可以減少向伺服器發送請求的次數，減輕伺服器的負載。

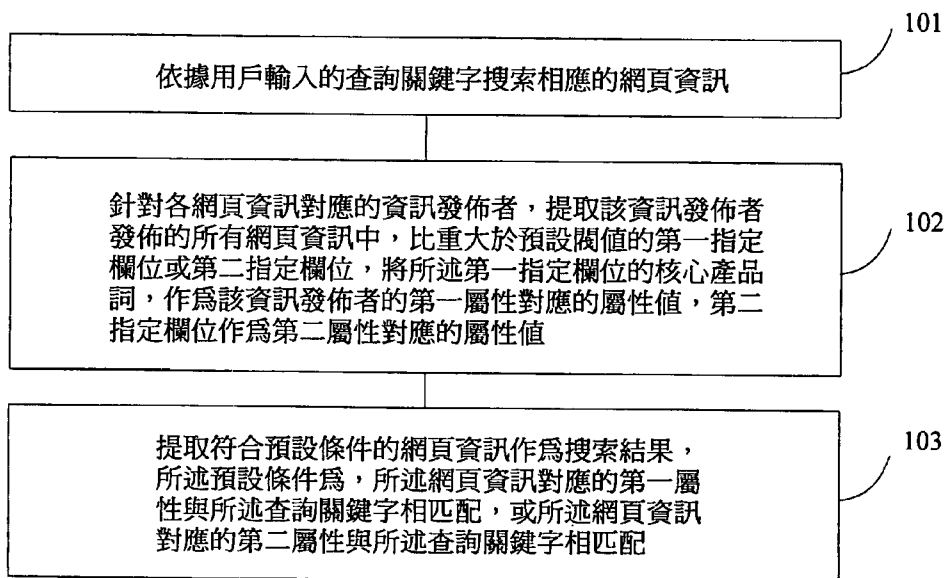


圖 1

發明專利說明書

(本申請書格式、順序，請勿任意更動，※記號部分請勿填寫)

※申請案號：101142220

※申請日：101年11月13日

※IPC分類：

G06F 17/30 (2006.01)

G06F 17/20 (2006.01)

G06Q 30/00 (2012.01)

一、發明名稱：(中文/英文)

搜索結果的產生方法及裝置

二、中文發明摘要：

本發明提供了一種搜索結果的產生方法和裝置，所述方法包括：依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。本發明可以減少向伺服器發送請求的次數，減輕伺服器的負載。

三、英文發明摘要：

四、指定代表圖：

(一) 本案指定代表圖為：第(1)圖。

(二) 本代表圖之元件代表符號簡單說明：無

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：無

六、發明說明：

【發明所屬之技術領域】

本發明涉及網路通信技術，特別是涉及一種搜索結果的產生方法和裝置。

【先前技術】

現有的搜索系統中，以電子商務網站搜索為例，當用戶在搜索頁面輸入搜索關鍵字進行檢索時，搜索引擎將查詢詞與供應資訊資料庫中的供應資訊進行匹配，對與查詢詞文本相似或語義相似的供應資訊，依據點擊率進行排序，將排名靠前的供應資訊返回給用戶。

以上現有技術中存在的問題是，由於搜索供應資訊的時候僅僅針對搜索關鍵字和供應資訊進行匹配，部分供應商為了增加商品曝光，會發佈一些虛假的供應資訊欺騙買家。例如某商家主營產品是鋼琴，為了吸引更多人購買鋼琴，就在網頁上投放了標題為供應某熱門手機的網頁資訊，買家點擊相應鏈結卻發現大量的關於鋼琴的網頁資訊，不符合買家的購買意圖，浪費了買家的搜索時間，進而增加了向伺服器發送請求的次數，加重了伺服器的負載。

因此，目前需要本領域技術人員解決的一個技術問題就是，提供一種搜索結果的產生機制，以減少向伺服器發送請求的次數，減輕伺服器的負載。

【發明內容】

本發明所要解決的技術問題是提供一種搜索結果的產生方法，以減少向伺服器發送請求的次數，減輕伺服器的負載。

本發明還提供了一種搜索結果的產生裝置，用以保證上述方法在實際中的應用及實現。

為了解決上述問題，本發明公開了一種搜索結果的產生方法，包括：

依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；

針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。

較佳的，所述第一屬性為主營產品，所述第二屬性為主營行業；

所述第一指定欄位為供應產品詞，所述第二指定欄位為所屬行業。

較佳的，所述網頁資訊為供應資訊時，所述網頁資訊包括供應產品詞、所屬行業、商品屬性、發佈者資訊，所

述發佈者資訊包括供應商地址、公司名稱、主營產品和主營行業。

較佳的，在針對各供應資訊，統計其供應商的主營產品或主營行業之前，所述方法還包括：

過濾與所述查詢關鍵字的相似度小於預設相似度的網頁資訊。

較佳的，所述網頁資訊與查詢關鍵字的相似度藉由以下步驟獲取：

分別提取各網頁資訊和查詢關鍵字中的核心產品詞；

由各核心產品詞在不同點擊類目中點擊率，計算各網頁資訊與查詢關鍵字的相似度。

較佳的，所述方法還包括：

將所述搜索結果進行排序後在頁面展示。

較佳的，所述方法還包括：

依據各資訊發佈者的第一屬性和第二屬性對其發佈者資訊進行更新。

較佳的，所述資訊發佈者為廣告主，所述將提取的網頁資訊進行排序的步驟包括：

分別統計所述查詢關鍵字在各網頁資訊中的特徵資訊，所述特徵資訊包括網頁資訊中查詢關鍵字的位置資訊，以及查詢關鍵字與網頁資訊的相關度；

由所述特徵資訊和點擊機率預測模型，獲取各網頁資訊的預測點擊機率，所述點擊機率預測模型包括不同的特徵資訊對應的預測點擊機率的統計值；

按照所述預測點擊機率和查詢關鍵字價格的乘積，對各網頁資訊進行排序。

本發明還提供了一種搜索結果的產生裝置，包括：

網頁資訊搜索模組，用於依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；

屬性值統計模組，用於針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

搜索結果提取模組，用於提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。

較佳的，所述裝置還包括：

網頁資訊過濾模組，用於過濾與所述查詢關鍵字的相似度小於預設相似度的網頁資訊。

與現有技術相比，本發明具有以下優點：

依據本發明，依據查詢關鍵字搜索供應資訊後，在過濾掉與查詢關鍵字相似度較低的搜索結果後，找出供應商的主營產品和主營行業，將主營產品包括查詢關鍵字對應的產品，或主營行業包括查詢關鍵字對應行業的供應資訊作為搜索結果，即搜索結果中的供應資訊均是以查詢關鍵字中的產品作為主營產品，或是主營行業的相關產品。若

某供應商發佈所查詢產品的虛假交易資訊，可以藉由找出其主營產品和主營行業並非所需查詢的產品，來過濾掉其供應資訊，增加了與買家的購買意圖的相關性，節約了用戶的搜索時間。並且，由於在搜索結果中剔除了與主營產品，或主營行業不相關的供應資訊，因此搜索引擎伺服器向發起搜索請求的用戶端傳輸搜索結果資料時，就相應的減少了傳輸的資料量（虛假交易資訊被過濾掉，傳輸的資料量必然減少），導致網路傳輸速度加快，避免網路堵塞。

此外，本發明還可以對各供應資訊的點擊機率進行預測，並依據預測點擊機率和查詢關鍵字的出價對供應資訊排序後進行展示，提高了搜索結果的點擊率。同時，依據各供應商的主營產品和主營行業對其供應商資訊進行更新，可以減少因為客戶更新不及時導致資訊遺漏的問題。

【實施方式】

為使本發明的上述目的、特徵和優點能夠更加明顯易懂，下面結合圖式和實施方式對本發明作進一步詳細的說明。

參考圖 1，示出了本發明的一種搜索結果的產生方法實施例 1 的流程圖，具體可以包括以下步驟：

步驟 101、依據用戶輸入的查詢關鍵字搜索相應的網頁資訊。

用戶在搜索頁面輸入查詢關鍵字進行檢索時，搜索引

引擎會將查詢關鍵字與預置的網頁資訊資料庫進行匹配，將與查詢關鍵字文本相似或語義相近的網頁資訊，提取出來作為初步搜索結果。在本發明中，所述網頁資訊可以為供應資訊。

步驟 102、針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值。其中，所述第一指定欄位的核心詞可以是核心產品詞。

其中，所述第一屬性為主營產品，所述第二屬性為主營行業，所述第一指定欄位為供應產品詞（例如可以是供應諾基亞手機），所述第二指定欄位為所屬行業（例如可以是通信設備），相應的，可以將第一指定欄位的核心產品詞作為第一屬性的屬性值（例如，第一指定欄位為“供應諾基亞手機”，提取的核心產品詞為“手機”，可以將手機作為主營產品的屬性值），第二指定欄位作為第二屬性的屬性值（例如，第二指定欄位為“通信設備”，將“通信設備”作為主營行業的屬性值）。在本發明中，所述網頁資訊為供應資訊時，可以包括供應產品詞、所屬行業、商品屬性、發佈者等資訊，其中發佈者資訊為供應商資訊。每條供應資訊可以對應相同的商品名稱，但商品屬性可以不一致。其中，所述發佈者資訊可以包括供應商地址、公司名稱、主營產品和主營行業等資訊。例如，用戶搜索手

機的供應資訊，其中一條供應資訊包括：

供應諾基亞手機

所屬行業：通信設備

GSM 紅色，產地廣州

供應商：User1，阿裏巴巴網路技術有限公司

主營產品：手機，MP3，電腦，電話機

主營行業：通信設備，商業服務。

在本發明的一種較佳實施例中，所述步驟 102 可以包括：

子步驟 S11、針對各網頁資訊對應的資訊發佈者，在資料庫中提取該資訊發佈者發佈的所有網頁資訊；

子步驟 S12、提取網頁資訊中的供應產品詞或所屬行業，計算各供應產品詞和所屬行業的比重；

子步驟 S13、針對比重大於預設閾值的供應產品詞，提取供應產品詞的核心產品詞作為該資訊發佈者的主營產品對應的屬性值，比重大於預設閾值的所屬行業作為該資訊發佈者的主營行業對應的屬性值。

針對各供應資訊，需要找出其主營產品或主營行業。首先，需要找出各供應資訊發佈者發佈的所有供應資訊，然後提取各供應資訊中的供應產品詞或所屬行業，然後在所有供應產品詞中提取比重較大的供應產品詞，將比重較大的供應產品詞的核心產品詞作為主營產品的屬性值，將比重較大的所屬行業作主營行業對應的屬性值。

在所有所屬行業中提取比重較大的所屬行業作為主營

行業對應的屬性值。具體的，可以將全量的供應資訊和供應商做關聯，提取出每個供應商對應的供應產品的集合，計算各個產品的比重，例如，第一預設閾值為 10%，某供應資訊發佈者的所有供應資訊中供應產品詞 A 出現了 50 次，供應產品詞 B 出現了 30 次，供應產品詞 C 出現了 2 次；供應產品詞 C 出現的比例為 $2/(50+30+2)$ 小於 10%，那供應產品詞 C 可能不是該供應資訊發佈者的主營產品對應的屬性值，而供應產品詞 A 和 B 出現的比例大於 10%，其對應的核心產品詞可以作為該供應資訊發佈者的主營產品對應的屬性值。

各供應產品詞含有一個核心產品詞，例如，諾基亞手機對應的核心產品詞為手機，具體的，可以使用通用的詞性標註的工具得到供應產品詞中每個詞的詞性（名詞、動詞、形容詞等），提取其中的名詞即為核心產品詞。

步驟 103、提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。

其中，本實施例所述第一屬性與查詢關鍵字相匹配可以是查詢關鍵字與第一屬性文字上相匹配、也可以是按照預先設定的對應關係上的相匹配。例如：第一屬性為“手機”，查詢關鍵字也是“手機”，則二者是文字上的相匹配；若第一屬性為“手機”，查詢關鍵字是“諾基亞”，若預先已建立“手機”與“諾基亞”之間的對應關係，則二者是按

照預先設定的對應關係上的相匹配。同理，第二屬性與查詢關鍵字相匹配可以是查詢關鍵字與第二屬性文字上相匹配、也可以是按照預先設定的對應關係上的相匹配。

針對搜索結果中的各供應資訊，找出其資訊發佈者的主營產品或主營行業之後，需要依據主營產品或主營行業對各供應資訊進行判斷，即需要對第一屬性或第二屬性不符合條件的搜索結果進行過濾。具體而言，即判斷用戶查詢關鍵字所指產品中的供應產品詞和各供應資訊的供應商資訊發佈者的主營產品是不是一致是否匹配，或是查詢關鍵字中產品的供應產品詞所對應的行業資訊和各供應資訊的資訊發佈者的主營行業是否匹配是不是一致，若其中任一條件判斷結果為是，則說明該網頁資訊和查詢關鍵字是相關的。若兩個條件的判斷結果均為否，則說明是虛假的網頁資訊，需要進一步過濾掉。如此，若某供應商發佈所查詢產品虛假交易資訊，可以藉由判斷其主營產品和主營行業均與查詢關鍵字不匹配，來過濾掉其供應資訊。

參考圖 2，示出了本發明的一種搜索結果的產生方法實施例 2 的流程圖，具體可以包括以下步驟：

步驟 201、依據用戶輸入的查詢關鍵字搜索相應的網頁資訊。

較佳的，本實施例在針對各網頁資訊統計其資訊發佈者的第一屬性和第二屬性之前，還可以對搜索得到的網頁資訊進行過濾，如圖 2 所示，所述方法還可以包括：

步驟 202、過濾與所述查詢關鍵字的相似度小於預設

相似度的網頁資訊。

具體的，所述網頁資訊與查詢關鍵字的相似度藉由以下步驟獲取：

子步驟 S21、分別提取各網頁資訊和查詢關鍵字中的核心產品詞；

子步驟 S22、由各核心產品詞在不同點擊類目中的點擊率，計算各網頁資訊與查詢關鍵字的相似度。

由於用戶填寫的查詢關鍵字沒有嚴格的標準，大多是口語類型的，不僅包含了用戶所想搜索的產品，還會包含了除產品之外的描述性詞語，並且描述方式多種多樣。例如用戶想要購買一款給老人使用的手機時，可能會在網頁上輸入以下查詢關鍵字“供應諾基亞手機，可供老人使用”或者“供應老人使用的多功能手機，諾基亞”，搜索的網頁資訊中可能會包含供老人使用的電腦，諾基亞顯示器等與用戶搜索關鍵字相關，但與用戶想要購買的產品不一致的網頁資訊。因此，需要對用戶查詢關鍵字的核​​心產品詞與搜索得到的網頁資訊的核心產品詞進行對比，計算出兩者的相似度，將相似度比較小的網頁資訊過濾掉。

具體的，首先提取出查詢關鍵字和網頁資訊中的核心產品詞，用戶輸入的查詢關鍵字和網頁資訊大體上是按照漢語的語法結構填寫的，用通用的詞性標註的工具能得到每個詞的詞性（名詞、動詞、形容詞等），提取其中的名詞即為核心產品詞。由於查詢關鍵字和網頁資訊中含有至少一個核心產品詞。如果只含有一個名詞，那麼其必為核

心產品詞；如果含有兩個名詞 A B，則可以判斷 A B 間的關係，若 A 修飾 B 的機率高，那麼 B 為核心產品詞，反之，A 為核心產品詞，或者機率差很低或者兩者是相似的名詞，則兩者均是核心產品詞；如果是大於等於 3 個名詞存在，則將最後一個名詞 C 作為核心產品詞，然後再判斷前面的各個名詞與 C 的關係（類似含有兩個產品詞的情況），將修飾機率低的 1 個或多個名詞作為最終的核心產品詞。

提取出查詢關鍵字和網頁資訊的核心產品詞之後，分別計算兩者的核心產品詞在不同的點擊類目中的點擊率，具體的，每條網頁資訊均攜帶了其所屬的類目資訊。當用戶搜索一個核心產品詞時，點擊供應資訊就相當於給這個核心產品詞添加了一個類目資訊，累積一定時間段後，統計該核心產品詞的所屬類目資訊，並統計整體的點擊類目分佈情況，得出核心產品詞在不同類目下的點擊率。

藉由核心產品詞的文本及其點擊類目分佈情況來判斷查詢關鍵字和網頁資訊中的核心產品詞的相似度，在獲取了兩者在相應的類目的點擊率後，可以藉由以下公式計算相似度：

$$D(Q_p, D_p) = \sum_w p_Q(w) \cdot \log \frac{p_Q(w)}{p_D(w)}$$

其中 Q_p 為查詢關鍵字（Query）的類目點擊分佈， D_p 為網頁資訊（Doc）中核心詞的類目點擊分佈， $p_Q(w)$ 為

Query 的類目 w 的點擊率， $p_D(w)$ 為 Doc 中核心詞的類目 w 的點擊率。

步驟 203、針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心產品詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值。

步驟 204、提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。

較佳的，在得到搜索結果後，還可以對所述搜索結果進行展示，如圖 2 所示，所述方法還可以包括步驟 205。

步驟 205、將所述搜索結果進行排序後在頁面展示。

將藉由兩次過濾的供應資訊置於一起再進行排序，最後展現給用戶。在本發明的一種較佳實施例中，若所述資訊發佈者為廣告主，所述步驟 205 可以包括：

子步驟 S31、分別統計所述查詢關鍵字在各網頁資訊中的特徵資訊，所述特徵資訊包括網頁資訊中查詢關鍵字的位置資訊，以及查詢關鍵字與網頁資訊的相關度；

子步驟 S32、由所述特徵資訊和點擊機率預測模型，獲取各網頁資訊的預測點擊機率，所述點擊機率預測模型包括不同的特徵資訊對應的預測點擊機率的統計值；

子步驟 S33、按照所述預測點擊機率和查詢關鍵字價

格的乘積，對各網頁資訊進行排序。

本發明不僅可以用於常用的網頁搜索結果的產生，還可以用於作為廣告的搜索結果產生。若網頁資訊是作為搜索結果進行展示，則可以依據各條網頁資訊的預測點擊機率進行排序；若網頁資訊是作為搜索廣告進行展示，則需要綜合考慮預測點擊機率和廣告主的出價情況。

預估點擊機率可以基於歷史點擊情況。記錄每次搜索的查詢關鍵字，當前展現的所有搜索結果及其在頁面上的位置，以及此次查詢下用戶點擊的網頁資訊。對於這些搜索，挖掘出查詢的關鍵字與被點擊的網頁資訊間的關係，即特徵資訊，如飄紅、命中詞的位置等位置資訊，以及查詢關鍵字與網頁資訊的相關度。

依據一段時間內搜索的多個資訊（可能有幾千萬的資料條），利用一些機器學習的方法（如邏輯迴歸模型），統計特徵資訊和預測點擊機率之間的關係，訓練出用戶在查詢一個關鍵字的情況下，什麼樣的特徵能夠得到什麼樣的點擊率，即建立點擊機率預測模型，其中包括了不同的特徵資訊對應的預測點擊機率的統計值。當新的查詢來的時候，即時計算即將展現的網頁資訊與查詢詞之間的關係，然後用這個模型來預測出點擊機率。

然後，分別統計所述查詢關鍵字在各網頁資訊中的特徵資訊，由所述特徵資訊和點擊機率預測模型，獲取各網頁資訊的預測點擊機率。供應商對於查詢詞的出價是從資料庫裏讀取的，用戶每買一個關鍵字都可以對它出價。最

後對網頁資訊進行排序是依據預測點擊率和出價的乘積。

較佳的，在得到搜索結果後，還可以對網頁發佈者的資訊進行展示，如圖 2 所示，所述方法還可以包括步驟 206。

步驟 206、依據各資訊發佈者的第一屬性和第二屬性對其發佈者資訊進行更新。

提取出主營產品後，若供應商自行填寫供應商資訊中的主營產品沒有包括該主營產品，則對其進行補充，對於供應商填寫的，除提取的主營產品之外的其他產品，則可以刪除。如上例，供應商自主填寫的主營產品是“手機、MP3、電腦、電話機”，提取其所有的供應資訊，發現該供應商除了該供應資訊之外，只有另一條 MP3 的供應資訊。依據這兩條供應資訊可以得知其主營產品為 MP3 和手機，因此，可以將供應商填寫的 MP3 和手機保留，將電腦和電話機刪除，使得供應商資訊更為準確，減少因為客戶更新不及時導致資訊遺漏的問題。

綜上所述，依據本發明，依據查詢關鍵字搜索供應資訊後，在過濾掉與查詢關鍵字相似度較低的搜索結果後，針對過濾後的供應資訊，找出供應商的主營產品和主營行業，將主營產品包括查詢關鍵字對應的產品，或主營行業包括查詢關鍵字對應行業的供應資訊作為搜索結果，即搜索結果中的供應資訊均是以查詢關鍵字中的產品作為主營產品，或是主營行業的相關產品。若某供應商發佈所查詢產品的虛假交易資訊，可以藉由找出其主營產品和主營行

業並非所需查詢的產品，來過濾掉其供應資訊，增加了與買家的購買意圖的相關性，節約了用戶的搜索時間。並且，由於在搜索結果中剔除了與主營產品，或主營行業不相關的供應資訊，因此搜索引擎伺服器向發起搜索請求的用戶端傳輸搜索結果資料時，就相應的減少了傳輸的資料量（虛假交易資訊被過濾掉，傳輸的資料量必然減少），導致網路傳輸速度加快，避免網路堵塞。

此外，本發明還可以對各供應資訊的點擊機率進行預測，並依據預測點擊機率和查詢關鍵字的出價對供應資訊排序後進行展示，提高了搜索結果的點擊率。同時，依據各供應商的主營產品和主營行業對其供應商資訊進行更新，可以減少因為客戶更新不及時導致資訊遺漏的問題。

對於方法實施例，為了簡單描述，故將其都表述為一系列的動作組合，但是本領域技術人員應該知悉，本發明並不受所描述的動作順序的限制，因為依據本發明，某些步驟可以採用其他順序或者同時進行。其次，本領域技術人員也應該知悉，說明書中所描述的實施例均屬於較佳實施例，所涉及的動作和模組並不一定是本發明所必須的。

參考圖 3，示出了本發明的一種搜索結果的產生裝置實施例 1 的結構方塊圖，具體可以包括以下模組：

網頁資訊搜索模組 301，用於依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；

屬性值統計模組 302，用於針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比

重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心產品詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

搜索結果提取模組 303，用於提取符合預設條件的網頁供應資訊作為搜索結果，所述預設條件為，所述網頁供應資訊對應的第一屬性與主營產品包括所述查詢關鍵字相匹配對應的產品，或所述網頁供應資訊對應的第二屬性與主營行業包括所述查詢關鍵字相匹配對應的產品所屬的行業。

在本發明的一種較佳實施例中，所述第一屬性可以為主營產品，所述第二屬性可以為主營行業；

所述第一指定欄位可以為供應產品詞，所述第二指定欄位可以為所屬行業。

在本發明的一種較佳實施例中，所述網頁資訊為供應資訊時，所述網頁資訊包括供應產品詞、所屬行業、商品屬性、發佈者資訊，所述發佈者資訊可以包括供應商地址、公司名稱、主營產品和主營行業。

在本發明的一種較佳實施例中，所述屬性值統計模組可以包括：

網頁資訊提取子模組，用於針對各網頁資訊所對應的資訊發佈者，在資料庫中提取該資訊發佈者發佈的所有網頁資訊；

比重計算子模組，用於提取網頁資訊中的供應產品詞

或所屬行業，計算各供應產品詞和所屬行業的比重。

屬性值提取子模組，用於針對比重大於預設閾值的供應產品詞，提取供應產品詞的核心產品詞作為該資訊發佈者的主營產品對應的屬性值，比重大於預設閾值的所屬行業作為該資訊發佈者的主營行業對應的屬性值。

參考圖 4，示出了本發明的一種搜索結果的產生裝置實施例 2 的結構方塊圖，具體可以包括以下模組：

網頁資訊搜索模組 401，用於依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；

較佳的，本發明在屬性值統計模組之前，可以先對搜索得到的網頁資訊進行過濾，如圖 4 所示，所述裝置還可以包括：

網頁資訊過濾模組 402，用於過濾與所述查詢關鍵字的相似度小於預設相似度的網頁資訊。

較佳的，所述網頁資訊與查詢關鍵字的相似度可以藉由以下子模組獲取：

核心產品詞提取子模組，用於分別提取各網頁資訊和查詢關鍵字中的核心產品詞；

相似度計算子模組，用於由各核心產品詞在不同點擊類目中點擊率，計算各網頁資訊與查詢關鍵字的相似度。

屬性值統計模組 403，用於針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將所述第一指定欄位的核心產品詞，作為該資訊發佈者的第一屬

性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

搜索結果提取模組 404，用於提取符合預設條件的網頁資訊作為搜索結果，所述預設條件為，所述網頁資訊對應的第一屬性與所述查詢關鍵字相匹配，或所述網頁資訊對應的第二屬性與所述查詢關鍵字相匹配。

較佳的，本發明在得到搜索結果之後，還可以對搜索結果進行排序，如圖 4 所示，所述裝置還可以包括：

搜索結果展示模組 405，用於將所述搜索結果進行排序後在頁面展示。

在本發明的一種較佳實施例中，所述資訊發佈者為廣告主，所述搜索結果展示模組可以包括：

特徵資訊統計子模組，用於分別統計所述查詢關鍵字在各網頁資訊中的特徵資訊，所述特徵資訊包括網頁資訊中查詢關鍵字的位置資訊，以及查詢關鍵字與網頁資訊的相關度；

點擊機率預測子模組，用於由所述特徵資訊和點擊機率預測模型，獲取各網頁資訊的預測點擊機率，所述點擊機率預測模型包括不同的特徵資訊對應的預測點擊機率的統計值；

供應資訊排序子模組，用於按照所述預測點擊機率和查詢關鍵字價格的乘積，對各網頁資訊進行排序。

較佳的，本發明還可以依據統計得到的供應商的主營產品和主營行業對其供應商資訊進行更新，如圖 4 所示，

所述裝置還可以包括：

資訊更新模組 406，用於依據各資訊發佈者的第一屬性和第二屬性對其發佈者資訊進行更新。

由於所述裝置實施例基本相應於前述圖 1 和圖 2 所示的方法實施例，故本實施例的描述中未詳盡之處，可以參見前述實施例中的相關說明，在此就不贅述了。

本發明可用於眾多通用或專用的計算系統環境或配置中。例如：個人電腦、伺服器電腦、手持設備或攜帶型設備、平板型設備、多處理器系統、基於微處理器的系統、機上盒、可編程的消費電子設備、網路 PC、小型電腦、大型電腦、包括以上任何系統或設備的分散式計算環境等等。

本發明可以在由電腦執行的電腦可執行指令的一般上下文中描述，例如程式模組。一般地，程式模組包括執行特定任務或實現特定抽象資料類型的常式、程式、物件、元件、資料結構等等。也可以在分散式計算環境中實踐本發明，在這些分散式計算環境中，由藉由通信網路而被連接的遠端處理設備來執行任務。在分散式計算環境中，程式模組可以位於包括儲存設備在內的本地和遠端電腦儲存媒體中。

最後，還需要說明的是，在本文中，術語“包括”、“包含”或者其任何其他變體意在涵蓋非排他性的包含，從而使得包括一系列要素的過程、方法、物品或者設備不僅包括那些要素，而且還包括沒有明確列出的其他要素，或

者是還包括為這種過程、方法、物品或者設備所固有的要素。在沒有更多限制的情況下，由語句“包括一個……”限定的要素，並不排除在包括所述要素的過程、方法、物品或者設備中還存在另外的相同要素。

以上對本發明所提供的一種搜索結果的產生方法，以及，一種搜索結果的產生裝置進行了詳細介紹，本文中應用了具體個例對本發明的原理及實施方式進行了闡述，以上實施例的說明只是用於幫助理解本發明的方法及其核心思想；同時，對於本領域的一般技術人員，依據本發明的思想，在具體實施方式及應用範圍上均會有改變之處，綜上所述，本說明書內容不應理解為對本發明的限制。

【圖式簡單說明】

圖 1 是本發明的一種搜索結果的產生方法實施例 1 的流程圖；

圖 2 是本發明的一種搜索結果的產生方法實施例 2 的流程圖；

圖 3 是本發明的一種搜索結果的產生裝置實施例 1 的結構方塊圖；

圖 4 是本發明的一種搜索結果的產生裝置實施例 2 的結構方塊圖。

【主要元件符號說明】

301：網頁資訊搜索模組

- 302：屬性值統計模組
- 303：搜索結果提取模組
- 401：網頁資訊搜索模組
- 402：網頁資訊過濾模組
- 403：屬性值統計模組
- 404：搜索結果提取模組
- 405：搜索結果展示模組
- 406：資訊更新模組

七、申請專利範圍：

1. 一種搜索結果的產生方法，其特徵在於，包括：
依據用戶輸入的查詢關鍵字搜索相應的網頁資訊；

針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將該第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

提取符合預設條件的網頁資訊作為搜索結果，該預設條件為，該網頁資訊對應的第一屬性與該查詢關鍵字相匹配，或該網頁資訊對應的第二屬性與該查詢關鍵字相匹配。

2. 如申請專利範圍第 1 項的方法，其中，該第一屬性為主營產品，該第二屬性為主營行業；

該第一指定欄位為供應產品詞，該第二指定欄位為所屬行業。

3. 如申請專利範圍第 2 項的方法，其中，該網頁資訊為供應資訊時，該網頁資訊包括供應產品詞、所屬行業、商品屬性、發佈者資訊，該發佈者資訊包括供應商地址、公司名稱、主營產品和主營行業。

4. 如申請專利範圍第 1 項的方法，其中，在針對各供應資訊，統計其供應商的主營產品或主營行業之前，該方法還包括：

過濾與該查詢關鍵字的相似度小於預設相似度的網頁

資訊。

5. 如申請專利範圍第 4 項的方法，其中，該網頁資訊與查詢關鍵字的相似度藉由以下步驟獲取：

分別提取各網頁資訊和查詢關鍵字中的核心產品詞；

由各核心產品詞在不同點擊類目中點擊率，計算各網頁資訊與查詢關鍵字的相似度。

6. 如申請專利範圍第 1 項的方法，其中，還包括：

將該搜索結果進行排序後在頁面展示。

7. 如申請專利範圍第 3 項的方法，其中，還包括：

依據各資訊發佈者的第一屬性和第二屬性對其發佈者資訊進行更新。

8. 如申請專利範圍第 6 項的方法，其中，該資訊發佈者為廣告主，該將提取的網頁資訊進行排序的步驟包括：

分別統計該查詢關鍵字在各網頁資訊中的特徵資訊，該特徵資訊包括網頁資訊中查詢關鍵字的位置資訊，以及查詢關鍵字與網頁資訊的相關度；

由該特徵資訊和點擊機率預測模型，獲取各網頁資訊的預測點擊機率，該點擊機率預測模型包括不同的特徵資訊對應的預測點擊機率的統計值；

按照該預測點擊機率和查詢關鍵字價格的乘積，對各網頁資訊進行排序。

9. 一種搜索結果的產生裝置，其特徵在於，包括：

網頁資訊搜索模組，用於依據用戶輸入的查詢關鍵字

搜索相應的網頁資訊；

屬性值統計模組，用於針對各網頁資訊對應的資訊發佈者，提取該資訊發佈者發佈的所有網頁資訊中，比重大於預設閾值的第一指定欄位或第二指定欄位，將該第一指定欄位的核心詞，作為該資訊發佈者的第一屬性對應的屬性值，第二指定欄位作為第二屬性對應的屬性值；

搜索結果提取模組，用於提取符合預設條件的網頁資訊作為搜索結果，該預設條件為，該網頁資訊對應的第一屬性與該查詢關鍵字相匹配，或該網頁資訊對應的第二屬性與該查詢關鍵字相匹配。

10. 如申請專利範圍第9項的裝置，其中，還包括：

網頁資訊過濾模組，用於過濾與該查詢關鍵字的相似度小於預設相似度的網頁資訊。

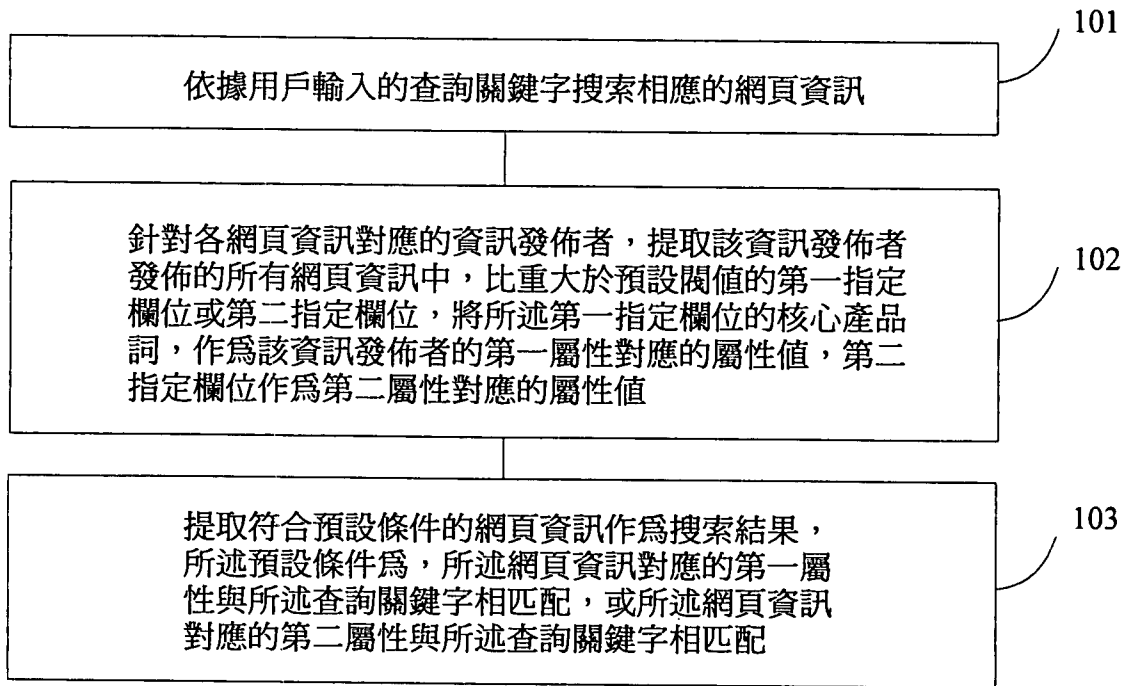


圖 1

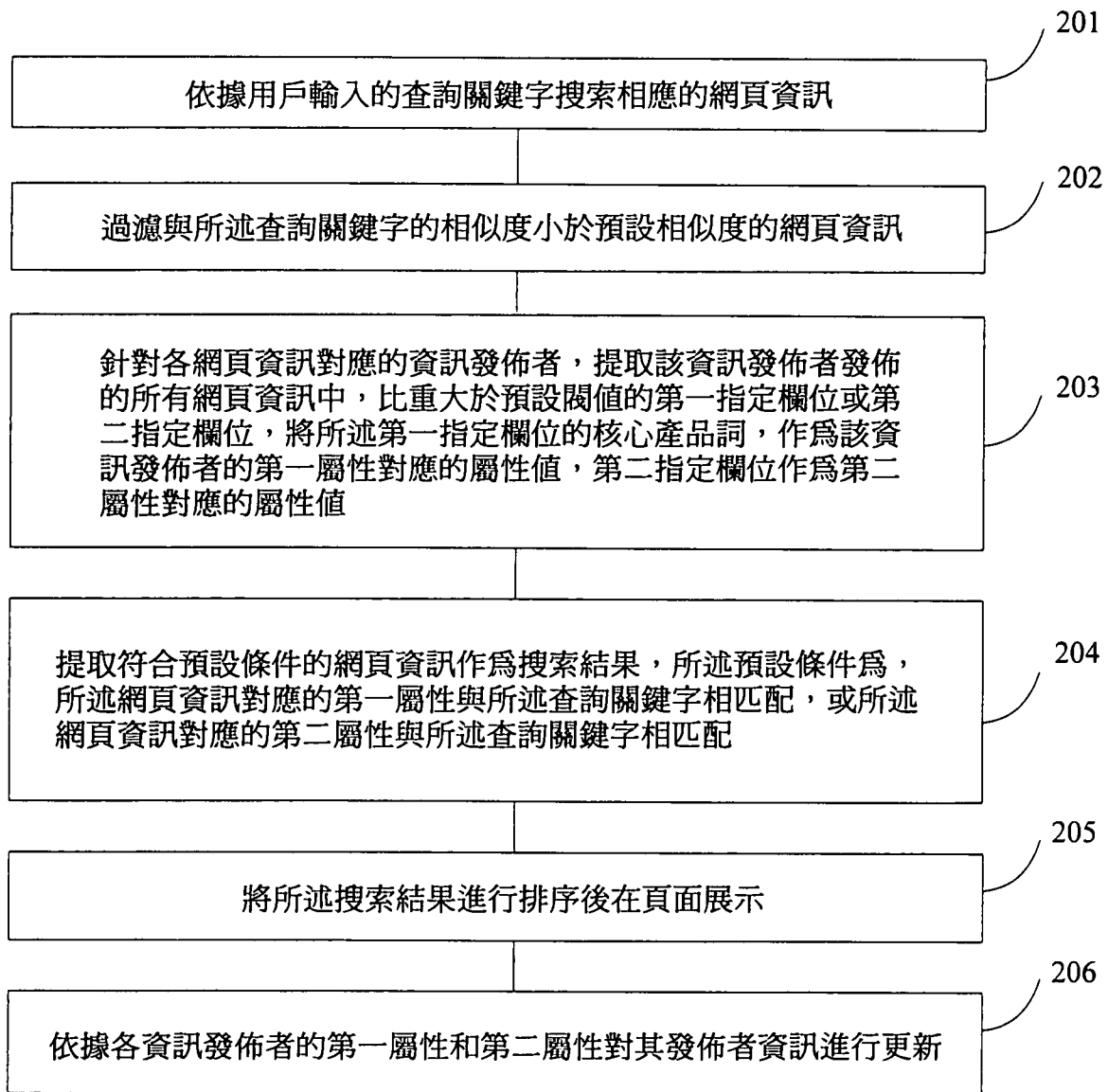


圖2

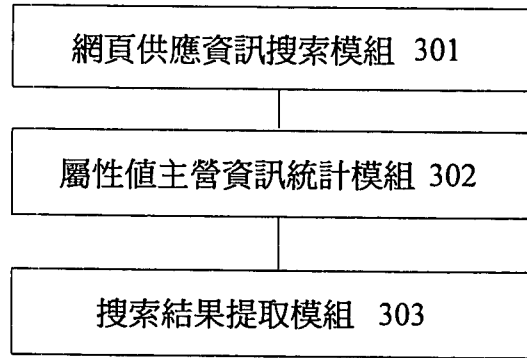


圖 3

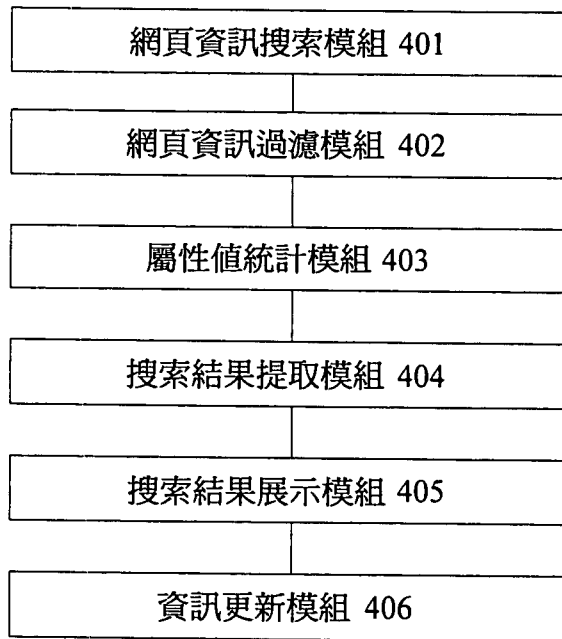


圖 4