



(21)申請案號：100143933

(22)申請日：中華民國 100 (2011) 年 11 月 30 日

(51)Int. Cl. : **G06Q30/02 (2012.01)**

(30)優先權：2011/08/16 中國大陸 201110235267.5

(71)申請人：阿里巴巴集團控股有限公司 (開曼群島) ALIBABA GROUP HOLDING LIMITED
(KY)

香港

(72)發明人：顧湘余 (CN)；傅其樂 (CN)；祖仲林 (CN)；寧偉 (CN)

(74)代理人：林志剛

(56)參考文獻：

TW 200639730A

TW 200945072A

US 2005/0222987A1

US 2009/0125719A1

US 2010/0332304A1

US 2011/0035379A1

審查人員：吳偉賢

申請專利範圍項數：15 項 圖式數：8 共 43 頁

(54)名稱

基於用戶行為的電子商務資訊推薦方法與裝置

(57)摘要

本申請提供了一種基於用戶行為的電子商務資訊推薦方法。該方法包括：統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率；判斷用戶行為類型數和用戶行為的頻率是否超過預設閾值，如果是，則將超過的資料過濾掉；根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；根據該用戶偏好資料進行內容資訊推薦。本申請還提供了一種基於用戶行為的電子商務資訊推薦裝置。本申請對用戶行為資料進行最佳化，剔除了可能存在的「偽資料」，使得在該資料基礎上分析得到的用戶偏好更加準確真實，處理速度得以提高，推薦效果更好。

指定代表圖：

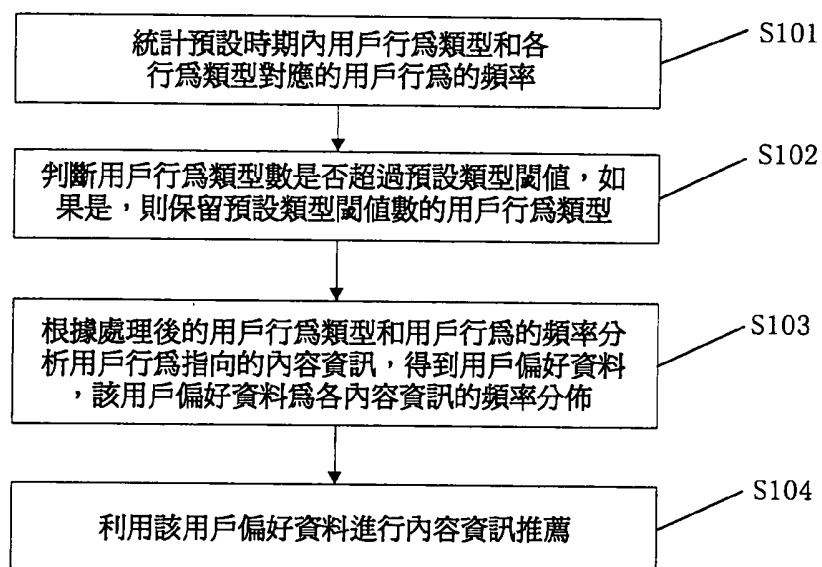


圖 1

六、發明說明：

【發明所屬之技術領域】

本申請係關於電子商務資訊推薦技術領域，尤其關於一種基於用戶行為的電子商務資訊推薦方法與裝置。

【先前技術】

隨著互聯網技術的發展，電子商務蓬勃興起。電子商務是利用電腦技術、網路技術和遠端通訊技術，使商務交易中的買賣雙方在不需謀面的情況下實現各種商貿活動。要實現商貿交易，買方必須預先獲知商品的內容資訊以便決策是否購買該商品。商品內容資訊包括商品所屬行業類目、商品的提供商、商品價格以及與該商品相關的資訊訊息等。在電子商務資訊交易平臺上，獲知該內容資訊的方式主要有兩種：一種是用戶（買方）主動搜尋方式，即用戶在明確自己的購買需求後，將需求以關鍵字的形式體現出來，然後利用該關鍵字在大量的多樣性資料中檢索，從而獲得需要的內容資訊；另一種是用戶被動接受型，即賣方透過電子商務交易平臺向用戶推薦商品的內容資訊，用戶被動接受賣方推薦的商品內容資訊後，在該內容資訊的引導下購買相關產品。對於第二種方式，商貿交易的賣方爲了提高交易成功率，往往不會將自己的所有資訊直接推薦給用戶，而是首先分析用戶過去的行為建立用戶偏好資料，然後根據用戶偏好針對性地推薦用戶可能感興趣的特定資訊。這種推薦方式在電子商務資訊交易中能顯著提高

用戶體驗，增加賣家曝光的精準度，能有效引導用戶迅速成為買家，從而降低交易成本。但是，電子商務資訊的資訊推薦者進行用戶偏好分析時獲取的用戶歷史行為資料可能存在偏差、錯誤，甚至包括以用戶名義惡意引導消費傾向的作弊資料，建立在這些資料基礎上的分析結果必然不能真實地反映用戶偏好特點，從而導致推薦內容資訊偏離用戶需求，推薦效果受到影響。另外，由於電子商務網站的資訊訪問量特別巨大，電子商務網站伺服器中存在海量的用戶行為資料，針對海量用戶行為資料分析得到用戶偏好資料對推薦系統自身處理能力是一種嚴峻考驗，海量的用戶行為資料減緩了推薦系統的分析處理速度。

【發明內容】

有鑑於此，本申請的發明目的在於提供一種基於用戶行為的電子商務資訊推薦方法與裝置，透過對用戶行為資料的最佳化，一方面濾除可能存在的「偽資料」以校正用戶偏好資料，從而確保根據用戶偏好資料推薦的內容資訊符合用戶的真實偏好；另一方面減少用於用戶偏好分析的資料量以提高偏好資料的分析速度，從而確保用戶快速的獲取到推薦的內容資訊。

本申請提供的基於用戶行為的電子商務資訊推薦方法包括：

統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率；

判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻率是否超過預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

利用用戶偏好資料向用戶推薦內容資訊。

本申請還提供了一種基於用戶行為的電子商務資訊推薦裝置，該裝置包括：

統計模組、判斷模組、偏好分析模組和資訊推薦模組，其中：

該統計模組，用於統計預設時期內的用戶行為類型和各行為類型對應的用戶行為的頻率；

該判斷模組，用於判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻率是否超過預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

該偏好分析模組，用於根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

該資訊推薦模組，用於利用用戶偏好資料向用戶推薦

內容資訊。

本申請的技術方案透過判斷用戶行為類型數與用戶行為的頻率是否符合預設閾值的方式對用戶的歷史行為資料進行最佳化，在該最佳化處理後的用戶行為資料基礎上分析得到用戶偏好，然後根據該偏好向用戶推薦相應的內容資訊。與現有技術相比，本申請的最佳化處理過程排除了用戶行為資料中的惡意行為、作弊行為等「偽資料」，使得根據該最佳化資料進行的偏好分析結果更真實地反映了用戶的偏好特點，進而使得根據該用戶偏好進行的資訊推薦的推薦效果更好，提高了電子商務資訊的交易成功率。另外，由於根據偏好分析結果推薦的資訊能夠符合用戶的實際需求願望，這樣用戶只需根據電子商務交易平臺推薦的資訊完成交易即可，而不必再反覆查找其所需的內容資訊，降低了反覆查找內容資訊給電子商務平臺帶來的訪問壓力。並且，由於本申請的最佳化處理過程去除了用戶行為資料中的大量「偽資料」，減少了用於資訊推薦分析的資料量，提高了資訊推薦分析的速度，能夠更加快速的獲取到推薦資訊。

【實施方式】

本申請透過統計用戶行為類型和用戶行為類型對應的用戶行為的頻率，將該統計結果與預設閾值進行比較，以濾除用戶行為資料中的惡意行為、作弊行為等「偽資料」，然後在上述經過最佳化處理後的用戶行為資料基礎上

進行用戶偏好分析，根據分析得到的用戶偏好特點實現相應內容資訊的推薦。

爲使本領域技術人員進一步瞭解本申請的特徵及技術內容，下面結合附圖和實施例，對本申請的技術方案進行詳細描述。

實施例一

● 電子商務資訊交易平臺是虛擬的市場，買賣主體不需要面對面磋商即可完成交易。交易的成功與否取決於交易中買方的求購願望與賣方的供給內容間的匹配程度。賣方主動提供商品資訊固然可以增加賣家的曝光率，提高交易成功機率，但是，如果賣方主動供給內容與買方需求偏差太大，賣方即便使出「渾身解數」，可能仍將甚微收穫。因此，爲提高交易成功率，賣家的主動推送行爲往往需要建立在對用戶的分析基礎上，這種分析需要預先收集大量的用戶資訊，包括用戶的歷史行爲資料、用戶屬性資料等內容，利用這些用戶訓練得到某個內容資訊的傾向預測資料，進而將某個特定用戶的用戶偏好與該傾向預測資料進行匹配以決定如何向用戶推薦相應的內容資訊。分析用戶偏好通常建立在用戶的歷史行爲資料基礎上，用戶歷史行爲反映了用戶的求購願望。比如，某個用戶行爲在一段時間內集中高頻出現，該行爲指向的內容資訊也集中出現在某些商品，則說明該用戶在該時期內的偏好就是這些商品，該用戶有極大可能會購買該商品，如果資訊提供商能夠

提供與該偏好匹配的產品資訊，必然增大了交易成功率。

然而，正如先前技術所言，有些用戶的歷史行為資料可能不真實，是「偽資料」，比如現實中可能存在這樣的情形：在某個時期內大部分用戶均關注某一產品，則說明該產品是該時段內的新潮產品，由於用戶的個人偏好大多受大眾消費趨勢的影響，那麼單個用戶的偏好將轉向該商品資訊，從而導致從事該商品交易的賣家利潤可觀。基於利潤的爭奪，有些賣家便抓住用戶的這種心理，以用戶名義在某時期內大量重複某些行為，以圖誘導消費趨勢。這樣用戶的歷史行為資料中將出現「作弊資料」，建立在這些資料基礎上的用戶偏好分析將「失真」，進而導致推薦的內容資訊與用戶的真實需求偏離，降低交易成功率。另外，由於分析用戶偏好需要建立在海量的用戶的歷史行為資料基礎上，再加上某些賣家惡意以用戶名義在某時期內大量重複某些行為，更進一步的增加了進行用戶偏好分析所依據的歷史行為資料量，導致用戶偏好的分析速度緩慢。本申請的實施例提供了一種基於用戶行為的電子商務資訊推薦方法，該方法對用戶歷史行為資料進行最佳化處理以濾除「偽資料」。參見附圖 1，本實施例提供的一種基於用戶行為的電子商務資訊推薦方法包括：

步驟 S101：統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率；

獲取用戶的偏好資料需要根據用戶的歷史行為進行分析，分析的前提是要選擇一個參考時間段，即設定一個預

設時期，該預設時期的長短將影響到用戶偏好的分析結果：預設時期長度不同，選擇的用戶行為類型的數量和用戶行為的頻率便不同，由此得出的用戶偏好也可能大不相同，一般而言，該預設時期不宜選擇過長，過長則無法準確反應用戶偏好的遷移特性和細微變化，也不宜選擇過短，過短則可能導致用戶的偏好不具有代表性；預設時期確定後，將統計分析該預設時期內的用戶行為類型以及該行為類型對應的用戶行為的頻率，這裏的用戶行為類型包括但不限於檢索、瀏覽、點擊、保存行為，用戶行為的頻率是某種用戶行為類型在該預設時期內發生的次數，用戶行為類型與用戶行為頻率對應；為後面更見形象的說明起見，此處假設在預設時期 T 內統計得到五種行為類型，分別為 $act1$ 、 $act2$ 、 $act3$ 、 $act4$ 、 $act5$ ，每種行為類型對應的用戶行為的頻率分別為 $f1$ 、 $f2$ 、 $f3$ 、 $f4$ 、 $f5$ ，由於「偽資料」的存在，上述五種用戶行為類型可能是虛假的，不代表用戶真實行為。

步驟 S102：判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留預設類型閾值數的用戶行為類型；

上面曾提到用戶行為可能包含有作弊行為，這些作弊行為在建立用戶行為偏好資料時需要予以剔除，否則將影響到根據偏好資料進行推薦的推薦效果；進行作弊資料剔除的依據是預設類型閾值，該預設閾值可以是透過分析一定時期內大量的用戶行為得出的統計資料，也可以是根據實際操作情況得出的一個經驗估計值，在此基礎上根據推

薦效果進行修正調整；上述舉例中，假設預設類型閾值 $\theta_{actth}=4$ ，那麼上述五種行為類型就超過了該預設閾值指定的數量，說明該用戶行為類型中包含有「偽行為」，這時則將上述的五種用戶行為類型按照預設類型閾值數進行保留，即只留下四種用戶行為類型，該四種用戶行為類型分別為 $act2$ 、 $act3$ 、 $act4$ 、 $act5$ ，由於用戶行為類型與用戶行為的頻率相聯繫，因而用戶行為的頻率也保留下 $f2$ 、 $f3$ 、 $f4$ 、 $f5$ 四個對應的頻率值，需要說明的是這裏過濾掉的是 $act1$ 行為類型，實際上，也可以過濾掉其他行為類型，只要剩餘的用戶行為類型數量不超過預設類型閾值即可；然而，儘管這種濾除操作並沒限定濾除物件，但在一些情況下結合其他因素進行濾除則更為合理，比如，在用戶行為類型數超過預設類型閾值時，優先考慮濾除時間最早的行為類型，因為該類型對用戶偏好的影響可能已經「過時」，也可以優先考慮將頻率最高的用戶行為對應的用戶行為類型濾除，因為該頻率最高的用戶行為可能恰巧代表的是作弊行為。該較佳實施例的推廣應用即為：在用戶行為類型數超過預設類型閾值時，保留低於預設頻率閾值的用戶行為對應的用戶行為類型，從而使得用戶行為類型數不超過預設類型閾值，透過這兩種方式可以有效濾除作弊行為。

步驟 S103：根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

用戶的每個行為均對應一個或者多個行為物件，該行為物件即是用戶行為指向的內容資訊，該內容資訊可以是單個資訊，也可以是複合資訊，比如用戶在網路中搜尋了一件商品，用戶行為類型則為「搜尋行為」，該「搜尋行為」指向的內容資訊即是該商品，該商品的屬性可以包括商品價格、商品形狀、商品所屬行業類目等中的一種，也可以是上述商品屬性中的幾種，甚至還可進一步延伸包括商品提供商、提供商的資質情況等；在上述的舉例中，假設上述每種用戶行為類型指向的內容資訊分別是： $act2$ 指向三種內容資訊 $C1$ 、 $C2$ 、 $C3$ ，該三種內容資訊出現的頻率分別為 f_{21} 、 f_{22} 、 f_{23} ，此三種頻率的總和為 f_2 ，表示為 $act2 (C1[f_{21}]、C2[f_{22}]、C3[f_{23}])$ ； $act3 (C1[f_{31}]、C2[f_{32}]、C3[f_{33}]、C4[f_{34}]、C5[f_{35}])$ 、 $act4 (C2[f_{42}]、C3[f_{43}]、C5[f_{45}])$ 、 $act5 (C2[f_{52}]、C3[f_{53}]、C4[f_{54}]、C5[f_{55}])$ ；得到上述資料後，以用戶行為指向的內容資訊為依據繪製用戶偏好資料曲線，將該曲線頂點（頻率值最大）對應的內容資訊作為用戶偏好：上述資料共出現 $C1$ 、 $C2$ 、 $C3$ 、 $C4$ 、 $C5$ 五種內容資訊，他們的頻率值分別為 $f_{21}+f_{31}$ 、 $f_{22}+f_{32}+f_{42}+f_{52}$ 、 $f_{23}+f_{33}+f_{43}+f_{53}$ 、 $f_{34}+f_{54}$ 、 $f_{35}+f_{45}+f_{55}$ ，選取其中頻率值最高的對應的內容資訊作為該用戶的用戶偏好資料，也可以先設定一個預設值，將超過該預設值的對應的多個內容資訊作為該用戶的用戶偏好資料。

步驟 S104：利用用戶偏好資料向用戶推薦內容資訊

。上述步驟中已經分析出了用戶的偏好資料，則根據該偏好資料即可進行內容推薦。

本申請的實施例透過將用戶的歷史行為類型和該類型對應的頻率與預設閾值進行比較，排除用戶歷史行為中的作弊資料、惡意資料，在該處理後的資料基礎之上進行用戶偏好分析，根據分析得到的用戶偏好資料實現內容資訊的推薦。與現有技術相比，本實例建立用戶偏好資料時對用戶行為資料進行了最佳化處理，消除了可能的「偽資料」，使得根據較為準確的用戶行為資料得到的用戶偏好能真實地反映用戶的實際需求願望，進而獲得較好的推薦效果，提高電子商務資訊交易的成功率。實際應用中，由於推薦的商品資訊能夠符合用戶的實際需求願望，這樣用戶只需根據電子商務交易平臺推薦的資訊完成交易即可，而不必再反覆查找其所需的商品資訊，降低了反覆查找商品資訊給電子商務平臺帶來的訪問壓力。另外，由於本申請的最佳化處理過程去除了用戶行為資料中的大量「偽資料」，減少了用於資訊推薦分析的資料量，提高了資訊推薦分析的速度，能夠更加快速的獲取到推薦資訊。

上述實施例中，僅僅從用戶行為類型的角度對用戶行為資料進行了最佳化，實際上，還可以對用戶行為的頻率進行最佳化處理，因為在預設時間段內某個用戶行為的頻率也可能出現異常，比如「畸高」或「畸低」，這些情形可能反映出該資料是作弊資料，也應當予以剔除。因

此，上實施例的步驟 S102 之後還可以增加下述步驟：

判斷各用戶行為的頻率是否超過該用戶行為的預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

經過上述步驟後的用戶行為類型和行為頻率更加得到最佳化，建立在此基礎上的用戶偏好更加準確進而取得更好的推薦效果。值得注意的一點是，上述步驟除可以和實施例一中步驟 S102 以並列方式進行外，還可以替代步驟 S102，同樣能達到改善推薦效果的目的。

實施例二

上述實施例雖然透過與預設類型（頻率）閾值的匹配濾除了用戶行為中的「偽資料」，經過濾除處理後的用戶行為資料已能較好地反映用戶的真實偏好。但是，在預設時期 T 內用戶行為的發生時間往往不同，發生時間的差異對分析用戶偏好的影響也不同，一般而言，越是靠近該預設時期 T 的結束時刻發生的用戶行為越接近用戶的真實偏好，越是靠近該預設時期 T 的起始時刻發生的用戶行為越偏離用戶的真實偏好。因此，有必要在上述實施例的基礎上對用戶行為資料進行時間影響消除處理，實現這種處理的具體方式很多，只要能反映出發生時間靠後的用戶行為對用戶偏好建立的影響大，發生時間靠前的用戶行為對用戶偏好建立的影響小，即可實現本申請的發明目的。在實施例一的基礎上增加上述處理步驟後的方法構成本申請

的又一個實施例，參見附圖 2，該實施例與實施例一除在步驟 S102、S103 之間添加下述內容外，其他步驟相同。

本實施例給出的時間影響消除處理的步驟包括：

對用戶行為進行時間影響消除處理，該處理步驟包括：
對用戶行為賦予時間權重值，該時間權重值為越靠近預設時期結束時刻權重值越高，越靠近預設時期的起始時刻權重值越低。

本實施例對用戶行為資料進行了時間影響消除處理，使得以用戶行為為基礎構建用戶偏好的準確度進一步提高，進而推薦效果得到更好改善。

儘管時間影響消除的具體實現方式較多，但本申請的實施例較佳地按照如下的方式計算時間權重值：

$$\text{時間權重值} = 1 / \log(D_{\text{off}})$$

公式中 D_{off} 為用戶行為發生時刻到預設時期結束時刻的時間差。

用戶行為被賦予權重值後，用戶行為的頻率值乘以相應的時間權重值以修正用戶行為頻率，按照這種方式最佳化後的用戶行為資料將更能準確地反映用戶在預設時期內的偏好特徵。

實施例三

用戶行為指向的內容資訊是分析用戶偏好的基礎，但是，用戶在進行某些用戶行為時指向的內容資訊受到用戶知識能力的限制，可能並不準確，由此可能導致在此基礎

上的分析出現偏差。比如，用戶實施了一個搜尋行為，該搜尋行為指向“關鍵字 1”，這個“關鍵字 1”是用戶根據自己掌握的知識所確定的關鍵字，但該關鍵字可能並不準確，甚至有歧義、錯誤，此刻則需進行「偏差校正」，校正後的資料才能用於進行用戶偏好分析。在實施例一的基礎上增加上述處理步驟後的方法可構成本申請的再一個實施例，參見附圖 3，該實施例與實施例一除在步驟 S102、S103 之間添加下述內容外，其他步驟相同。本實施例給出的內容資訊偏差校正方法包括：

對內容資訊進行標準化處理，該處理步驟包括：根據標準產品單元和行業知識庫校正內容資訊。

標準產品單元 SPU (Standard Product Unit) 是一組可複用、易檢索的標準化資訊的集合，該集合描述了「產品」的特性，儲存了產品的基本資訊；行業知識庫是某行業的「知識」的集合，這些「知識」包括行業標準、行業業務模型、行業業務資料與資訊模型、行業資訊化全景圖、行業領域構件、行業採購鏈、行業上下游產品等內容，該知識庫涵蓋的內容相當豐富。用戶行為指向的內容資訊提取出來後，將其與標準產品單元和行業知識庫進行匹配比較，如果兩者不一致，則用標準產品單元和行業知識庫中的標準語言描述該內容資訊，經過這樣標準化處理後的內容資訊更加便於後續步驟的識別和處理，對於快速、準確地完成用戶偏好分析具有良好效果。

實施例四

前述實施例選取了一個參考時間週期即預設時間 T ，透過對該時期內的用戶行為類型和用戶行為的頻率進行分析得到用戶偏好資料。但是，用戶在該時間週期內可能發生偏好遷移（變化），也就是說，用戶可能在一個時段內對內容資訊 1 感興趣，在另一個時段內可能對與內容資訊 1 相關聯的內容資訊 2 感興趣（內容資訊 1 與內容資訊 2 形成遷移對），這種情形即是用戶偏好的遷移特性，在分析具有偏好遷移特性的用戶行為時，需要注意的是：只有遷移對的遷移終點對應的內容資訊（內容資訊 2）才真實地代表用戶當下的興趣，對於遷移對的遷移起點對應的內容資訊（內容資訊 1）表示用戶已經「放棄」該內容資訊，對此不再感興趣，因此，應當將遷移終點對應的內容資訊作為用戶偏好，並據此進行內容資訊的推薦。除用戶偏好發生遷移的情形外，在預設時期內的某個時段內可能存在用戶對某個內容資訊極度感興趣，用戶行為頻率極大，這種情形構成用戶偏好的突發特性，該特性說明該用戶的偏好並不穩定，僅代表該用戶的激情偏好，而不能代表整個預設時期的偏好，因此也應當在分析用戶偏好時予以剔除。與用戶偏好的突發特性相對應的是用戶偏好的持續特性，即在一個時段內用戶的偏好集中在某幾個內容資訊上，在另一個時段內用戶偏好仍然集中在該幾個內容資訊上，這種情形說明該用戶的偏好較為穩定，這些用戶行為資料則能較好地反映用戶偏好，根據此類用戶偏好進行的

內容資訊推薦有極大的可能符合該用戶的需求，進而提高交易成功率。由此，本申請給出了又一個實施例，該實施例中增加將上述突發性、遷移性用戶偏好過濾掉、保留下持續性用戶偏好。本實施例除將實施例一的步驟 S103 變化為下述內容外，其他步驟相同。參見附圖 4，本實施例給出的分析得到用戶偏好資料的步驟包括：

步驟 A：根據用戶行為發生的時間將處理後的用戶行為類型和用戶行為頻率在預設時期內分成至少兩個等長時段單元；

預設時期選取的時段長度越短，用戶偏好的細微程度越大，透過比較不同時段的用戶偏好曲線，即可得到突發性用戶偏好和遷移性偏好；將預設時期 T 劃分成至少兩個等長時段單元，根據實際的需要可選擇兩個或多個時段單元，時段單元越多，得到的用戶偏好變化情，也越細微，這裏為方便起見，以兩個時段單元為例： $T1$ 、 $T2$ ， $T1+T2=T$ 。

步驟 B：根據各時段單元內的用戶行為類型和用戶行為頻率分析用戶行為指向的內容資訊，得到該時段單元的用戶偏好資料；

劃分時段單元後，分別分析 $T1$ 、 $T2$ 時段單元的用戶行為類型和用戶行為頻率分析用戶行為指向的內容資訊，分析過程可以採取繪製曲線的方式進行：以用戶內容資訊為橫軸、用戶內容資訊出現的頻率為縱軸。

步驟 C：將各時段單元的用戶偏好資料進行比對以濾

除遷移性用戶偏好資料對中的遷移起點資料和/或突發性用戶偏好資料，得到整個預設時期的用戶偏好資料；

得到 T1、T2 這兩個時段的用戶偏好曲線圖後，將兩個曲線圖進行比對分析，將兩個時段中均具有的內容資訊作為持續性偏好資料予以保留，將僅在前一個時段單元中出現的內容資訊作為遷移性用戶偏好的遷移起點予以濾除，將僅在後一個時段單元中出現的內容資訊作為遷移性用戶偏好的遷移終點予以保留，將在任何一個時段單元中出現的突發性用戶偏好資料予以濾除，進而得到整個預設時段 T 的用戶偏好資料。

本實施例透過將預設時期進行細分，然後比對細分後的每個時段單元的用戶偏好資料的變化，濾除了遷移性、突發性用戶偏好資料。這樣得到的用戶偏好更為準確地反映用戶在整個預設時期內的用戶偏好，進而在此用戶偏好資料基礎上進行的資訊內容推薦的推薦效果更好。

實施例五

上述實施例為得到較為準確的用戶行為資料，進行了一系列的最佳化操作，濾除了用戶行為中的「偽資料」，為分析得到準確的用戶偏好奠定了較好的基礎。但是，對於用戶行為指向的內容資訊是否準確還待研究，因為分析用戶偏好的落腳點是內容資訊，而這些內容資訊由電子商務資訊賣家提供，賣家為爭奪市場利益、搶佔用戶，也會惡意地在內容資訊上製造「偽資料」，比如，當內容資

訊包含產品價格資訊時，賣家可能惡意標定低價以打擊對手，或者在壟斷情況下隨意標定高價以攫取更多利潤，這些行為擾亂了電子商務資訊交易市場，基於這些「偽」的內容資訊進行用戶偏好分析得到的結果可能大為「失真」。為此，有必要對內容資訊進行「打假」。在實施例一的基礎上增加上述處理步驟後的方法可構成本申請的又一個實施例，參見附圖 5 本實施例與實施例一除在步驟 S102、S103 之間添加下述內容外，其他步驟相同。本實施例給出的排除上述「偽資料」的方法包括：

判斷產品標定價格是否在預設價格範圍內，如果否，則將產品價格設置在預設價格範圍內；和/或，判斷產品的所屬行業與提供該產品的供應商的註冊行業是否一致，如果否，則將該產品資訊濾除。

經過本實施例對內容資訊的最佳化後，進行用戶偏好分析的資料源更加準確，有利於得到更加有效的用戶偏好資料。

實施例六

上述幾個實施例該的方法在得到用戶的偏好資料後，資訊推薦步驟的具體實現方法可以採用現有技術。現有技術通常是將內容資訊獨立和分別地進行推薦，這種推薦效果並不好。為此，本申請的實施例提供了一種協同推薦的方法，該方法透過對大量的用戶內容資訊分析得到內容資訊的群體偏好資料，然後根據用戶偏好資料和群體偏好資

料進行內容資訊推薦。本實施例除將實施例一的步驟 S104 變換為下述內容外，其他步驟相同。參見附圖 6，本實施例給出的內容資訊推薦的步驟包括：

步驟 A'：選取預設數量的其他用戶組成參考用戶集，計算該用戶與參考用戶集內各用戶的相似度，將相似度超過預設相似度閾值的用戶組成相似用戶集；

根據用戶的歷史行為分析得到用戶偏好，進而依據該用戶偏好向用戶推薦內容資訊，固然迎合了用戶的求購願望，實現了提高交易成功率的目的，但是，實際上每個用戶的用戶偏好都不會侷限在以往的用戶偏好上，他的偏好是不斷擴展變化的，而這種擴展變化往往受限於他自身的能力和獲得資訊的管道，導致不能得到較好的擴展，此種情況下，與該用戶具有相似或相同用戶偏好的其他用戶的用戶偏好資料則值得借鑒，因為可以推知他們除具有上述相同的用戶偏好外，還可能在其他方面也具有相同的偏好；基於這種分析，則需要選擇一定數量的用戶組成參考用戶集，計算參考用戶集中的用戶與該用戶的相似度，對相似度超過預設閾值的用戶則構成該用戶的相似用戶集，相似用戶集內的用戶的偏好資料對該用戶具有較高參考價值；需要說明的是計算用戶之間的相似度可以採用多種方法，比如聚類演算法，該方法在現有技術中已有大量應用，此處不再贅述。

步驟 B'：分析相似用戶集內各用戶的用戶行為得到該相似用戶集的群體偏好資料，該群體偏好資料為用戶行為

指向的內容資訊的頻率分佈；

透過步驟 A'獲得相似用戶集後，分析相似用戶集的各用戶的用戶行為得到群體偏好資料，具體的分析方法和前述實施例的方式相同，群體偏好資料是與單個用戶的用戶偏好資料「同質」的概念，只是資料量的不同在資料結果上存在差異，均表示用戶行為指向的內容資訊的頻率分佈情況。

步驟 C'：根據用戶偏好資料和該群體偏好資料向用戶推薦內容資訊。

本實施例透過計算相似度，找出與該用戶具有相似特徵的其他用戶，並將其他用戶的用戶行為資料作為對該用戶進行內容資訊推薦的參考，較好地擴展了向用戶推薦的內容資訊，增強了用戶體驗，進一步拓寬了電子商務資訊交易的領域和範圍。

本實施例在計算得到用戶的相似用戶集時，對「其他用戶」的選取並沒有作特別限定，實際上，由於地理環境的差異、各地風土人情的不同，同一個地域範圍內用戶的用戶偏好資料可能更具有參考價值，因此，本實施例較佳地按照下面的方式選擇參考用戶集。

判斷該用戶 IP 地址對應的地域與該用戶註冊的地域是否一致，如果是，則選取該用戶註冊地域地的其他用戶組成參考用戶集；如果否，則選取該用戶 IP 地址對應的地域地的其他用戶組成參考用戶集。

透過以用戶登錄地的用戶資料作為分析群體偏好的基

礎，不僅由於同一地域的用戶行為資料相比其他地方的用戶行為資料的可參考性更大，使得分析得到的用戶偏好結果更準確，據此進行推薦的推薦效果更好，極大提高了電子商務資訊的交易成功率，而且由於採用就近原則，極大方便了用戶在電子商務資訊交易完成後儘快獲得相應的產品或服務。

實施例七

在實施例六中已經提到得到用戶的偏好資料後，資訊推薦步驟的具體實現方法可以採用現有技術，也可以採用實施例六該的協同推薦方法，協同推薦方法將待推薦用戶的自身偏好資料與一定範圍內的群體偏好資料結合起來，擴展了向待推薦用戶推薦的內容資訊，這是一種「橫向」擴展，實際上，除橫向擴展方式，本申請還提供一種縱向擴展方式的實施例，即將用戶偏好資料中的多個用戶偏好與預設規則集進行匹配，將匹配成功的某個或某些偏好作為向用戶推薦內容資訊的依據，這種方法稱為聯合推薦方法。本實施例除將實施例一的步驟 S104 變換為下述內容外，其他步驟相同。參見附圖 7，本實施例給出的內容資訊推薦的步驟包括：

步驟 A”：選取預設數量的用戶組成規則創建用戶集，分析該規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料，根據各用戶的用戶偏好資料訓練得到規則集，該規則集為用以預測內容資訊推薦的各類用戶偏好

的組合規則集；

透過分析用戶的歷史行為得到的用戶偏好資料，大多數情況下這種用戶偏好是複合偏好，也就是說該用戶偏好資料中包含了該用戶在各個層面、各個領域、各的時段的多個偏好，比如產品偏好、供應商偏好、產品詞偏好等，向用戶進行內容資訊推薦時儘管可以按照這些所有的偏好向用戶推薦，但是，這種「廣撒網」的方式在某些情況下並不能達到提高交易成功率的目的，反而容易導致用戶在眾多的內容資訊中無所選擇，因此，採用少而精、高命中率的方式向用戶推薦內容資訊將是一種可取的方式；爲了從用戶的眾多偏好中找出「有效」的用戶偏好，需要首先選取一定數量的用戶組成規則創建用戶集並根據這些用戶的用戶偏好建立規則集，進而利用該規則集作爲選取用戶偏好的依據；規則創建用戶集內的各用戶偏好資料的獲取可以採用現有技術的方式得到，也可以按照本申請所提到方法對用戶行為資料最佳化後分析得到；得到各用戶的用戶偏好資料後，將用戶的各個偏好與該偏好對應的頻率組成該用戶的向量特徵組，然後對所有的向量特徵組進行訓練得到規則集，具體的訓練方式現有技術已公開，這裏不再贅述。

步驟 B”：將待推薦用戶的用戶偏好資料與該規則集進行匹配，根據匹配成功的偏好資料向待推薦用戶推薦內容資訊。

獲得到規則集後，將待推薦用戶的用戶偏好與規則集

中的規則進行匹配，如果匹配成功則將根據相應的用戶偏好向該用戶推薦內容資訊。

爲了更加清楚的說明上述過程，這裏舉例闡釋：假設選取的規則創建用戶集爲 100 個，每個用戶的偏好包括產品、供應商、產品詞、求購、資訊等五個偏好，透過分析 100 個用戶的歷史行爲資料得到了各個用戶的偏好資料，偏好資料包括內容資訊和該內容資訊的頻率，爲了便於比較將頻率進行歸一化處理，該歸一化的基準可以是 10 次；透過按照現有技術的方法訓練 100 個用戶的用戶偏好資料得到三條規則，現在需要向 A1、B2、C3 三個用戶進行推薦，推薦時將 A1 的用戶偏好資料與規則集內的規則進行匹配，然後按照匹配成功的用戶偏好推薦內容資訊。上述過程請參見如下表格：

用戶名稱 用戶偏好類別		A1	B2	C3	規則集
產品偏好	歸一化頻率 偏好指向類目	0.6 數學	0 無	0.2 數學	規則 1：如果產品偏好的歸一化頻率高於 0.5，則推薦產品偏好對應內容資訊；
供應商偏好	歸一化頻率 偏好指向類目	0.2 英語	0 無	0.1 數學	
產品詞偏好	歸一化頻率 偏好指向類目	0.3 數學	0 無	1 語文	
求購偏好	歸一化頻率 偏好指向類目	0 無	1 語文	0 無	規則 2：如果產品詞偏好的歸一化頻率高於產品偏好與資訊偏好歸一化頻率之和的 1.2 倍，則推薦產品詞偏好對應內容資訊；
資訊偏好	歸一化頻率 偏好指向類目	0 無	0.5 英語	0.5 英語	
群體偏好	歸一化頻率 偏好指向類目	0 無	4 英語	0 無	
匹配規則集結果		符合規則 1	符合規則 3	符合規則 2	規則 3：如果群體偏好的歸一化頻率高於資訊品號與求購偏好歸一化頻率之和的 2 倍，則推薦群體偏好對應內容資訊。
推薦內容資訊		產品偏好 數 學	群體偏好 英 語	產品詞偏好 語文	

值得注意的是上述表格中列出了群體偏好一欄，前面介紹了根據規則創建用戶集的各用戶的用戶行為分析得到了用戶偏好資料，實際上，也可以根據該集合中的用戶偏好資料進一步得到這個群體的群體偏好，進而利用該群體偏好來創建關於群體偏好的規則。

本實施例將用戶的多種偏好結合起來，選擇最最佳化的推薦方案進行內容資訊推薦，避免了傳統推薦方法根據所有的偏好毫無取捨的「全盤」推薦，增強了推薦效果，有效提高了交易成功率。

實施例八

上述七個實施例均是本申請的方法實施例，對應地，本申請還提供了裝置實施例。參見附圖 8，本申請的實施例提供的一種基於用戶行為的電子商務資訊推薦裝置 800 包括：

統計模組 801、判斷模組 802、偏好分析模組 803 和資訊推薦模組 804，其中：

統計模組 801，用於統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率；

判斷模組 802，用於判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻率是否超過預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

偏好分析模組 803，用於根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

804 資訊推薦模組，用於利用用戶偏好資料向用戶推薦內容資訊。

本裝置的工作過程是：統計模組 801 統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率，並將資料傳遞給判斷模組 802，由該模組判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻

率是否超過該用戶行為的預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；然後由偏好分析模組 803 根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，並將該資料傳遞給資訊推薦模組 804，由該模組利用用戶偏好資料進行內容資訊推薦。

本申請的技術方案在根據用戶偏好進行資訊推薦之前，透過判斷用戶行為類型數與用戶行為頻率是否符合預設值的方式對用戶的歷史行為進行最佳化，最佳化後的用戶行為資料排除了惡意行為和作弊行為，在該用戶行為資料基礎上分析得到的用戶偏好資料較為真實地反映了用戶的偏好特點，進而使得根據該用戶偏好資料進行資訊推薦的推薦效果更好，有利於促進電子商務資訊交易的正常發展。此外，由於推薦的商品資訊能夠符合用戶的實際需求願望，這樣用戶只需根據電子商務交易平臺推薦的資訊完成交易即可，而不必再反覆查找其所需的商品資訊，降低了反覆查找商品資訊給電子商務平臺帶來的訪問壓力。並且，由於本申請的最佳化處理過程去除了用戶行為資料中的大量「偽資料」，減少了用於資訊推薦分析的資料量，提高了資訊推薦分析的速度，能夠更加快速的獲取到推薦資訊。

本實施例的裝置 800 還可以包括時間影響消除模組 805，用於在分析用戶行為指向的內容資訊前，對用戶行為賦予時間權重值，該時間權重值為越靠近預設時期結束

時刻權重值越高，越靠近預設時期起始時刻權重值越低。該模組消除時間影響後有利於更加準確地分析用戶行為資料。

本實施例的裝置 800 偏好分析模組 803 可以有多種組成機構，只要能實現本申請的發明目的即可，但是，本申請的實施例優選按照如下結構方式組成該模組，該模組包括：

時段單元劃分子模組 8031、偏好分析子模組 8032 和過濾子模組 8033，其中：

時段單元劃分子模組 8031，用於根據用戶行為發生的時間將處理後的用戶行為類型和用戶行為的頻率在預設時期內劃分成至少兩個時段單元；

偏好分析子模組 8032，用於根據各時段單元內的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到該時段單元的用戶偏好資料；

過濾子模組 8033，用於根據各時段單元的用戶偏好資料過濾掉遷移性、突發性偏好資料以得到預設時期的用戶偏好資料。

本實施例的裝置 800 還可以包括相似度計算模組 806，用於在選取預設數量的其他用戶組成參考用戶集後，計算待推薦用戶與參考用戶集內各用戶的相似度，將相似度超過預設相似度閾值的用戶組成相似用戶集；則：偏好分析模組 803 還用於分析相似用戶集內各用戶的用戶行為得到該相似用戶集的群體偏好資料；資訊推薦模組 804，還

用於根據待推薦用戶的用戶偏好資料和該群體偏好資料向該用戶推薦內容資訊。

本實施例的裝置 800 還包括規則集創建模組 807，用於在選取預設數量的用戶組成規則創建用戶集，調用偏好分析模組分析該規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料，根據各用戶的用戶偏好資料訓練得到規則集，該規則集為用以預測內容資訊推薦的各類用戶偏好的組合規則集，則：偏好分析模組還用於分析規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料，該用戶偏好資料為用戶行為指向的內容資訊的頻率分佈；資訊推薦模組，還用於將待推薦用戶的用戶偏好資料與該規則集進行匹配，根據匹配成功的用戶偏好資料向待推薦用戶推薦內容資訊。包含規則集創建單元 807 的裝置將用戶的多種偏好結合起來，選擇最佳化的推薦方案進行內容資訊推薦，避免了傳統推薦方法根據所有的偏好毫無取捨的「全盤」推薦，增強了推薦效果，有效提高了交易成功率。

以上所述僅為本申請的較佳實施例，並不用以限制本申請，凡在本申請的精神和原則之內，所作的任何修改、等同替換、改進等，均應包含在發明的保護範圍之內。

【圖式簡單說明】

圖 1 為本申請的方法的實施例一的流程圖；

圖 2 為本申請的方法的實施例二的流程圖；

圖 3 爲本申請的方法的實施例三的流程圖；

圖 4 爲本申請的方法的實施例四的流程圖；

圖 5 爲本申請的方法的實施例五的流程圖；

圖 6 爲本申請的方法的實施例六的內容資訊推薦的流程圖；

圖 7 爲本申請的方法的實施例七的內容資訊推薦的流程圖；

圖 8 爲本申請的裝置的實施例八的組成框圖。

【主要元件符號說明】

801：統計模組

802：判斷模組

803：偏好分析模組

8031：時段單元劃分子模組

8032：偏好分析子模組

8033：過濾子模組

804：資訊推薦模組

805：時間影響消除模組

806：相似度計算模組

807：規則集創建模組

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：100143933

※申請日：100 年 11 月 30 日

※IPC 分類：G06Q 3/02 (2012.01)

一、發明名稱：(中文／英文)

基於用戶行為的電子商務資訊推薦方法與裝置

二、中文發明摘要：

本申請提供了一種基於用戶行為的電子商務資訊推薦方法。該方法包括：統計預設時期內用戶行為類型和各行為類型對應的用戶行為的頻率；判斷用戶行為類型數和用戶行為的頻率是否超過預設閾值，如果是，則將超過的資料過濾掉；根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；根據該用戶偏好資料進行內容資訊推薦。本申請還提供了一種基於用戶行為的電子商務資訊推薦裝置。本申請對用戶行為資料進行最佳化，剔除了可能存在的「偽資料」，使得在該資料基礎上分析得到的用戶偏好更加準確真實，處理速度得以提高，推薦效果更好。

三、英文發明摘要：

七、申請專利範圍：

1. 一種基於用戶行為的電子商務資訊推薦方法，其特徵在於，該方法包括：

統計預設時期內的用戶行為類型和各行為類型對應的用戶行為的頻率；

判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻率是否超過預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

利用該用戶偏好資料向用戶推薦內容資訊。

2. 根據申請專利範圍第 1 項所述的方法，其中，在分析用戶行為指向的內容資訊前，該方法還包括：對用戶行為進行時間影響消除處理，該處理步驟包括：對用戶行為賦予時間權重值，該時間權重值為越靠近預設時期結束時刻權重值越高，越靠近預設時期的起始時刻權重值越低。

3. 根據申請專利範圍第 2 項所述的方法，其中，該時間權重值按照下述公式得到：

$$\text{時間權重值} = 1 / \log(D_{\text{off}})$$

公式中 D_{off} 為用戶行為發生時刻到預設時期結束時刻的時間差。

4. 根據申請專利範圍第 1 項所述的方法，其中，在分析用戶行為指向的內容資訊前，該方法還包括：對內容資訊進行標準化處理，該處理步驟包括：根據標準產品單元和產業知識庫校正內容資訊。

5. 根據申請專利範圍第 1 項所述的方法，其中，該分析用戶行為指向的內容資訊得到用戶偏好資料的步驟包括：

根據用戶行為的發生時間將處理後的用戶行為類型和用戶行為的頻率在預設時期內分成至少兩個時段單元；

根據各時段單元內的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到該時段單元的用戶偏好資料；

根據各時段單元的用戶偏好資料過濾掉遷移性、突發性偏好資料以得到整個預設時期的用戶偏好資料。

6. 根據申請專利範圍第 1 項所述的方法，其中，該內容資訊包括產品資訊時，則該分析用戶行為指向的內容資訊得到用戶偏好資料的步驟包括：

判斷產品標定價格是否在預設價格範圍內，如果否，則將產品價格設置在預設價格範圍內；和/或，判斷產品的所屬產業與提供該產品的供應商的註冊產業是否一致，如果否，則將該產品資訊濾除。

7. 根據申請專利範圍第 1 項所述的方法，其中，利用用戶偏好資料向用戶推薦內容資訊包括：

選取預設數量的其他用戶組成參考用戶集，計算待推

薦用戶與參考用戶集內各用戶的相似度，將相似度超過預設相似度閾值的用戶組成相似用戶集；

分析相似用戶集內各用戶的用戶行為得到該相似用戶集的群體偏好資料，該群體偏好資料為用戶行為指向的內容資訊的頻率分佈；

根據待推薦用戶的用戶偏好資料和該群體偏好資料向待推薦用戶推薦內容資訊。

8. 根據申請專利範圍第 7 項所述的方法，其中，該選取預設數量的其他用戶組成參考用戶集包括：

判斷待推薦用戶的 IP 地址對應的地域與該用戶註冊的地域是否一致，如果是，則選取待推薦用戶註冊地域地的其他用戶組成參考用戶集；如果否，則選取待推薦用戶的 IP 位址對應的地域地的其他用戶組成參考用戶集。

9. 根據申請專利範圍第 1 項所述的方法，其中，利用用戶偏好資料向用戶推薦內容資訊包括：

選取預設數量的用戶組成規則創建用戶集，分析該規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料，根據各用戶的用戶偏好資料訓練得到規則集，該規則集為用以預測內容資訊推薦的各類用戶偏好的組合規則集；

將待推薦用戶的用戶偏好資料與該規則集進行匹配，根據匹配成功的偏好資料向待推薦用戶推薦內容資訊。

10. 根據申請專利範圍第 1 項所述的方法，其中，當用戶行為類型數超過預設類型閾值時，保留的用戶行為類

型包括：頻率小於等於預設頻率閾值的用戶行為對應的用戶行為類型。

11. 一種基於用戶行為的電子商務資訊推薦裝置，其特徵在於，該裝置包括：

統計模組、判斷模組、偏好分析模組和資訊推薦模組，其中：

該統計模組，用於統計預設時期內的用戶行為類型和各行為類型對應的用戶行為的頻率；

該判斷模組，用於判斷用戶行為類型數是否超過預設類型閾值，如果是，則保留小於等於預設類型閾值數的用戶行為類型；和/或，判斷各用戶行為的頻率是否超過預設頻率閾值，如果是，則將預設頻率閾值作為該用戶行為的頻率；

該偏好分析模組，用於根據上述處理後的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到用戶偏好資料，該用戶偏好資料為各內容資訊的頻率分佈；

該資訊推薦模組，用於利用用戶偏好資料向用戶推薦內容資訊。

12. 根據申請專利範圍第 11 項所述的裝置，其中，該裝置還包括時間影響消除模組，用於在分析用戶行為指向的內容資訊前，對用戶行為賦予時間權重值，該時間權重值為越靠近預設時期結束時刻權重值越高，越靠近預設時期起始時刻權重值越低。

13. 根據申請專利範圍第 11 項所述的裝置，其中，該偏好分析模組包括：

時段單元劃分子模組、偏好分析子模組和過濾子模組，其中：

該時段單元劃分子模組，用於根據用戶行為的發生時間將處理後的用戶行為類型和用戶行為的頻率在預設時期內劃分成至少兩個時段單元；

該偏好分析子模組，用於根據各時段單元內的用戶行為類型和用戶行為的頻率分析用戶行為指向的內容資訊，得到該時段單元的用戶偏好資料；

該過濾子模組，用於根據各時段單元的用戶偏好資料過濾掉遷移性、突發性偏好資料以得到整個預設時期的用戶偏好資料。

14. 根據申請專利範圍第 11 項所述的裝置，其中，該裝置還包括相似度計算模組，用於在選取預設數量的其他用戶組成參考用戶集後，計算待推薦用戶與參考用戶集內各用戶的相似度，將相似度超過預設相似度閾值的用戶組成相似用戶集；則：

該偏好分析模組還用於分析相似用戶集內各用戶的用戶行為得到該相似用戶集的群體偏好資料，該群體偏好資料為用戶行為指向的內容資訊的頻率分佈；

該資訊推薦模組，還用於根據待推薦用戶的用戶偏好資料和該群體偏好資料向待推薦用戶推薦內容資訊。

15. 根據申請專利範圍第 11 項所述的裝置，其中，

該裝置還包括規則集創建模組，用於在選取預設數量的用戶組成規則創建用戶集，調用偏好分析模組分析該規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料，根據各用戶的用戶偏好資料訓練得到規則集，該規則集為用以預測內容資訊推薦的各類用戶偏好的組合規則集，則：

該偏好分析模組還用於分析規則創建用戶集內各用戶的用戶行為得到各用戶的用戶偏好資料；

該資訊推薦模組，還用於將待推薦用戶的用戶偏好資料與該規則集進行匹配，根據匹配成功的用戶偏好資料向待推薦用戶推薦內容資訊。

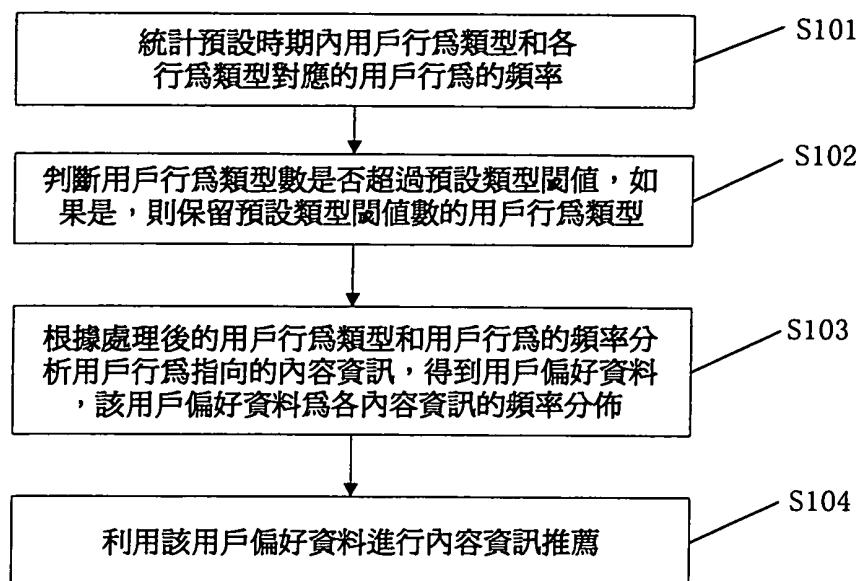


圖 1

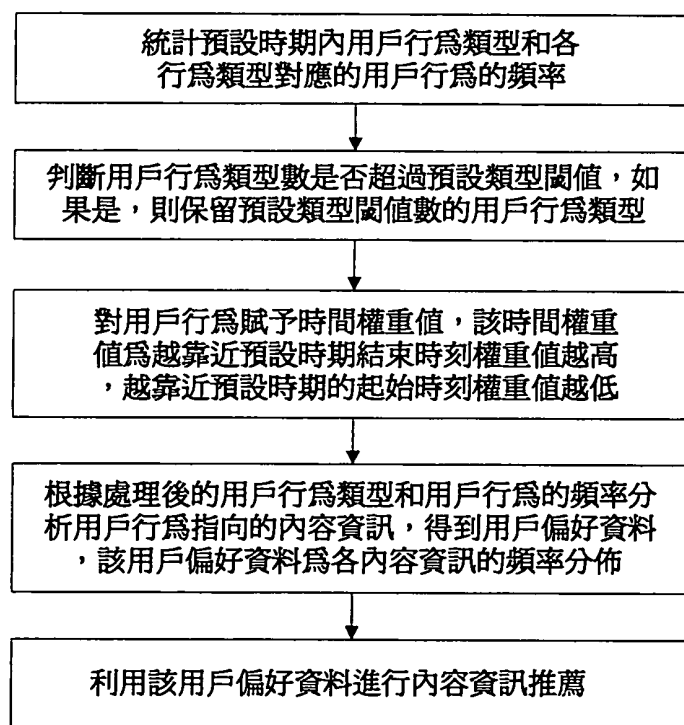


圖 2

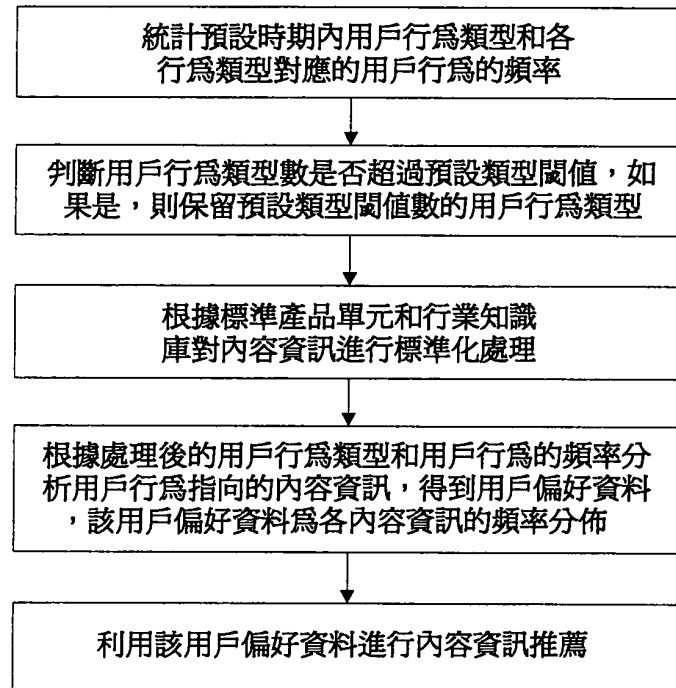


圖 3

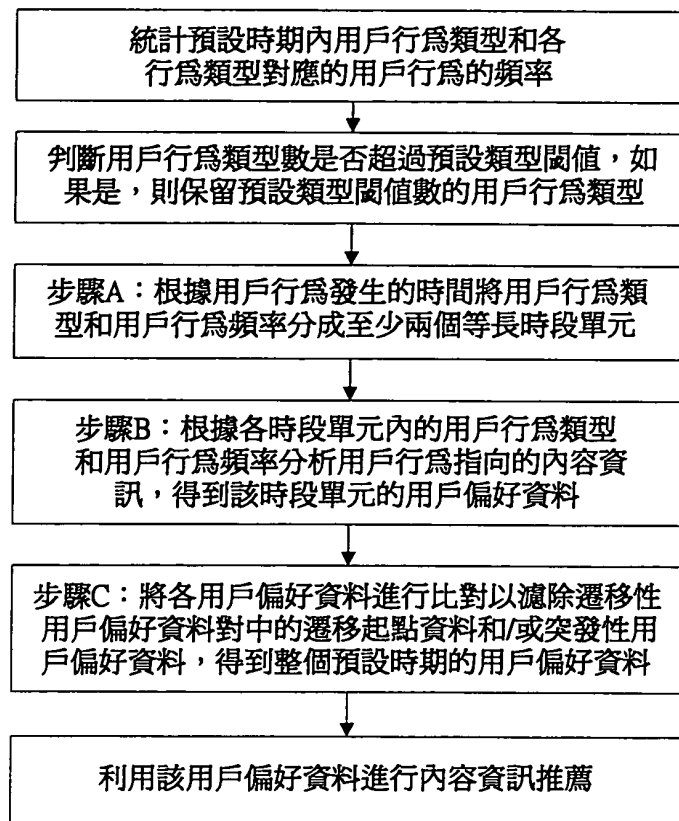


圖 4

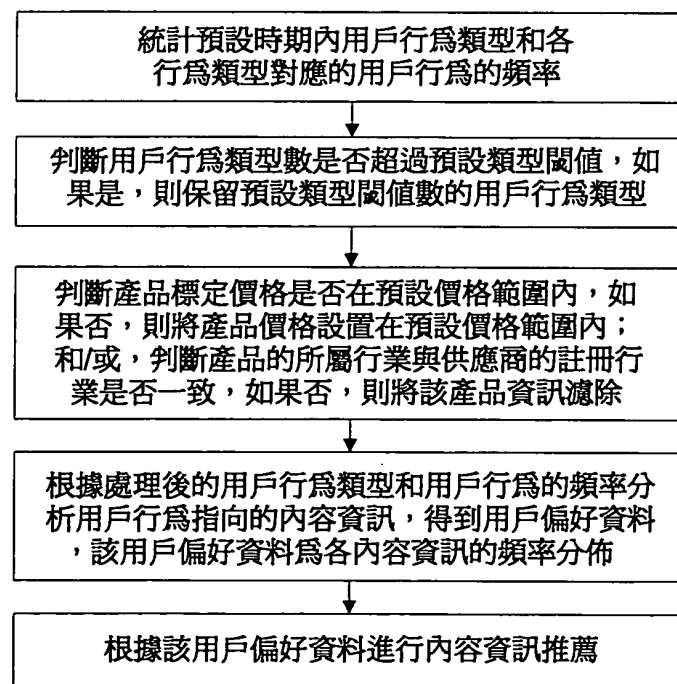


圖 5

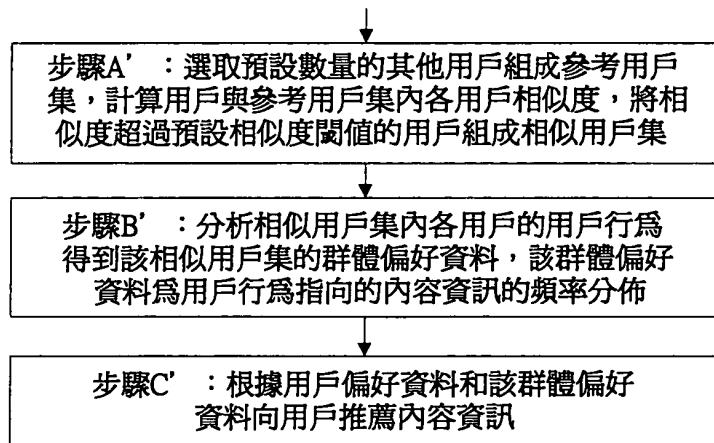


圖 6

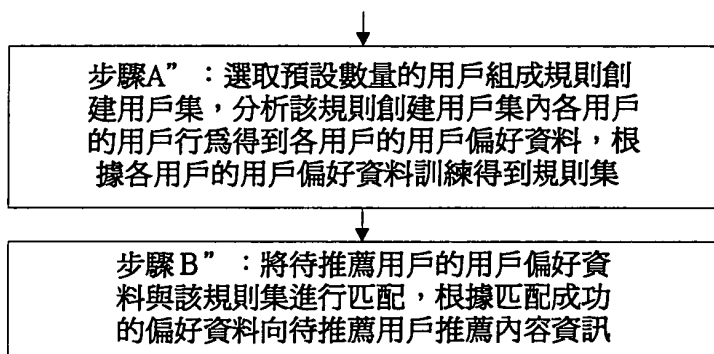


圖 7

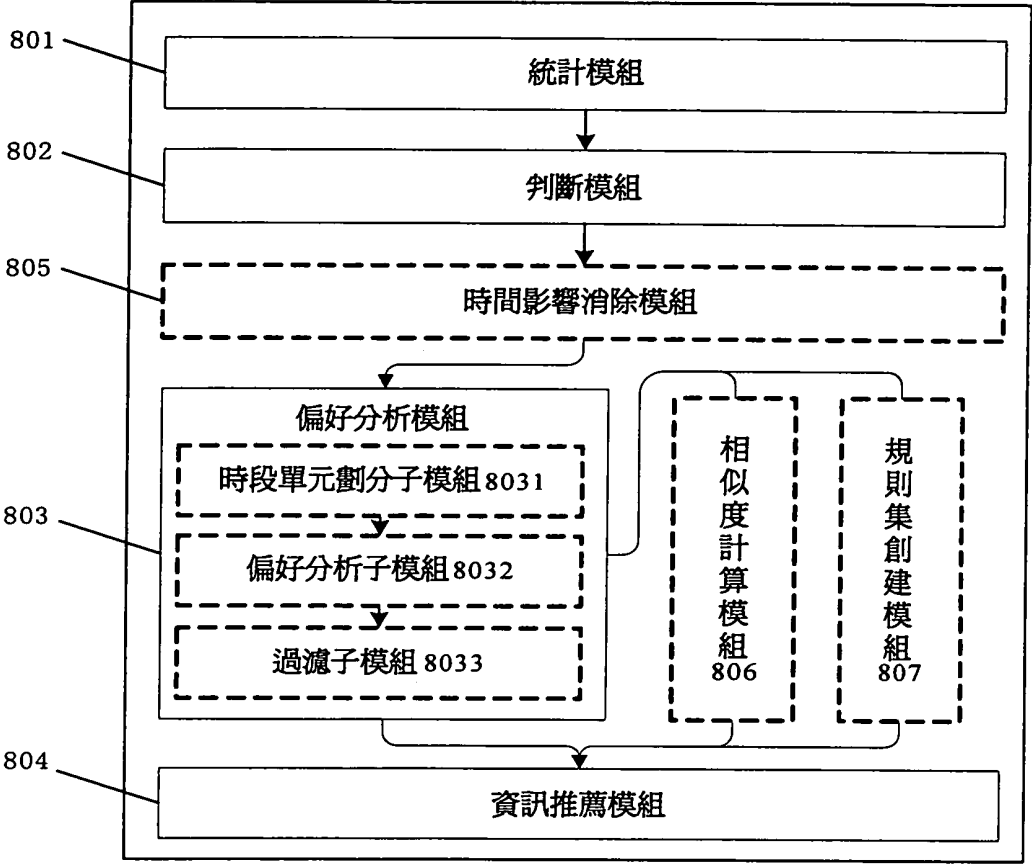


圖 8

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