



發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：92128546

※ 申請日期：92-10-15 ※IPC 分類：G06F 19/00

壹、發明名稱：(中文/英文)

線上分析處理之方法及系統

METHOD AND SYSTEM FOR ONLINE ANALYTICAL PROCESSING
(OLAP)

貳、申請人：(共 1 人)

姓名或名稱：(中文/英文)

美商萬國商業機器公司

INTERNATIONAL BUSINESS MACHINES CORPORATION

代表人：(中文/英文)

傑拉德 羅森賽

GERALD ROSENTHAL

住居所或營業所地址：(中文/英文)

美國紐約州阿蒙市新果園路

NEW ORCHARD ROAD, ARMONK, NY 10504 U.S.A.

國 籍：(中文/英文)

美國 U.S.A.

參、發明人：(共 2 人)

姓 名：(中文/英文)

1. 派崔克 亞拉斯
PATRICK ARRAS
2. 亞方斯 約翰尼 史汀霍夫
ALFONS JOHANN STEINHOFF

住居所地址：(中文/英文)

1. 德國威爾德史坦市辛特爾街13號
HINTERE STR. 13, D-71263 WEIL DER STADT, FEDERAL
REPUBLIC OF GERMANY
2. 德國亨利堡市克奇哈德區5號
KIRCHHALDE 5, D-71083 HERRENBERG, FEDERAL REPUBLIC
OF GERMANY

國 籍：(中文/英文)

- 1.-2. 均德國 GERMANY

肆、聲明事項：

☐ 本案係符合專利法第二十條第一項 ☐ 第一款但書或 ☐ 第二款但書規定之期間，其日期為： 年 月 日。

☒ 本案申請前已向下列國家（地區）申請專利：

1. 歐洲專利機構；2002年10月18日；02023362.3
- 2.
- 3.
- 4.
- 5.

☒ 主張國際優先權(專利法第二十四條)：

【格式請依：受理國家（地區）；申請日；申請案號數 順序註記】

1. 歐洲專利機構；2002年10月18日；02023362.3
- 2.
- 3.
- 4.
- 5.

☐ 主張國內優先權(專利法第二十五條之一)：

【格式請依：申請日；申請案號數 順序註記】

- 1.
- 2.

☐ 主張專利法第二十六條微生物：

☐ 國內微生物 【格式請依：寄存機構；日期；號碼 順序註記】

☐ 國外微生物 【格式請依：寄存國名；機構；日期；號碼 順序註記】

☐ 熟習該項技術者易於獲得，不須寄存。

玖、發明說明：

【發明所屬之技術領域】

本發明通常係關於電腦所執行的資料庫管理系統，尤其係關於對應的OLAP(線上分析處理)與資料倉儲應用。更具體的說，本發明係關於一種用於產生資料記錄之使用者定義的呈現(presentation)或檢視(view)之方法及系統，該等資料記錄被包含在該資料庫管理系統中。更具體的說，本發明係關於具有大量資料記錄的資料庫管理系統。

【先前技術】

資料倉儲與線上分析處理(OLAP)應用已突顯對快速儲存及擷取多維資料的需求。資料倉儲係一企業的多種商務系統所收集的資料的所有部分或重要部分之中心儲存庫(repository)。有時術語"資訊倉儲"被使用來取代資料倉儲。通常，資料倉儲被容納在企業主要伺服器上。在一資料倉儲資料庫上有選擇地擷取且組織來自多種線上交易處理(OLTP)應用及其他來源之資料，以供分析應用與使用者查詢使用。資料倉儲作業強調自不同來源擷取資料，以進行有用的分析及存取，但是通常不是就一般使用者或可能需要存取專業(有時為區域)資料庫的知識工作者的觀點出發。後一種技術通常稱為資料超市，資料採擷(data mining)、資訊網採擷及決策支援系統(DSS)是可利用資料倉儲之三種應用。

OLAP係讓使用者能夠輕鬆且有選擇地自不同觀點擷取並檢視資料的電腦處理。舉例而言，使用者可請求：分析

資料，並以試算表顯示某家電信公司在該年度中特定月份在某國家銷售之所有的行動電話產品，比較該特定月份與先前月份之相同產品的營收數字，接著觀看在相同期間裏該國家中其他產品銷售的比較。

為促進此類分析，會將OLAP資料儲存在一種多維資料庫中。儘管一關聯性資料庫可視為兩維資料庫，但是一種多維資料庫將每一資料屬性(通常稱為"資料索引鍵"，例如產品、地理銷售區域和期間)視為個別獨立的維度。OLAP軟體可定位多個維度之交集，例如，「在某時期中，在東部區域以高於某個價格而出售之所有產品」，並顯示該等維度。多種屬性(例如時期)可細分為子屬性。

類似於資料倉儲，OLAP可被用於資料採擷或發現資料項之間先前未察覺到的關係。由於不是所有的異動資料都為趨勢分析所需要，所以OLAP資料庫不需要像資料倉儲一樣大。藉由開放式資料庫連接(ODBC)，可自現有的關聯性資料庫匯入資料，以建置一適用於OLAP的多維資料庫。

如上所述，主要是按照支援一般使用者分析及導覽活動之合併(consolidated)資料的動態多維分析來確定OLAP功能性，包括：

整個維度(dimensions)、階層架構(hierarchies)及/或成員所套用的計算及模型化；

- 跨越連續期間的趨勢分析；
- 為在螢幕上進行檢視而分割(slice)子集；
- 鑽研至合併(consolidation)之較深層級；

- 直接到達基礎詳細資料(detail data)；及
- 轉換成在檢視區域中之新的維度比較。

通常會在一多使用者用戶端/伺服器環境中建構OLAP，且不管資料庫大小及複雜性，都試圖對資料庫存取提供穩定快速的回應。資料庫存取所傳送之資料通常會藉由如下方式而呈現給一般使用者：交叉表(cross-table)；樞紐分析呈現(pivot presentations)；或主要由已呈現之資料記錄單元格排列定義之基礎資料庫的部分甚至整體內容的檢視。在使用一普通電腦螢幕的通常狀況下，該排列必須是兩維的。樞紐分析呈現之另一實例係在一電腦記憶體、資料庫或其類似物中提供的二維資料陣列。

該基礎資料庫之每一資料記錄由許多資料屬性或資料索引鍵組成，如圖1所示，其中由資料索引鍵的數目來確定該等資料記錄之大小，且因同樣地由資料索引鍵的數目來決定此整個資料庫之大小。

一特定樞紐分析呈現通常取決於資料記錄(如顧客資訊、賣給顧客之對應物品，及銷售每一物品的時間)內容之特殊索引鍵排列的預指定使用者喜好設定(preference)。因此，經常需要在該樞紐分析呈現之單元格中，彙總不同資料記錄的內容或給定的資料記錄的資料索引鍵元素。

存在許多用於在一多維資料庫中產生資料記錄之樞紐分析呈現的方法。第一種方法使用著名的簡單查詢語言(SQL)，來藉由資料庫查詢而獲得一現有的資料庫之使用者定義的樞紐分析檢視。但是因為必須存取該資料庫中儲存

之所有資料記錄，以檢查其是否符合該等基礎查詢條件，所以在一較大型資料庫中之樞紐分析檢視的計算要求相當多的處理資源。

就第二種方法而言，其具有已知的OLAP軟體工具，其中因為藉由超立方(hyper-cubic)資料結構在內部處理資料記錄，所以需要大儲存容積以儲存中間產生的OLAP資訊。

根據第三種方法，結合上述第一種與第二種方法，以提供一對應的混合概念。但是，因為效能顯著變化，並且高度依賴實際的樞紐分析呈現，所以該第三種方法也是不利的。

【發明內容】

因此本發明之一目標係提供一種電腦所建構之方法及系統，該方法及系統用於產生一資料庫管理系統中包含之資料記錄的使用者定義之 n 維矩陣呈現或檢視，該等呈現或檢視更具效率，並因執行速度比相對應先前技術的方法更快，且避免了上述先前技術方法的上述缺點。

另一目標係提供一種可運用在一具有合適回應時間之"線上分析處理"(OLAP)環境中的相對應方法及系統。

還有一目標係提供一種僅需要最少量之處理資源來產生該等矩陣呈現或檢視的相對應方法及系統。

以上目標可由獨立的申請專利範圍之特點獲得。有利的實施例為該等子申請專利範圍(subclaim)的主題(subject matter)。

本發明之基本觀點係多維資料庫之基礎資料記錄的序列

化，例如該等基礎資料記錄以一連續資料序列的形式排列成一 m 維事實資料表。該序列因此由替代符號值 (placeholder value) (較佳為整數) 組成，其中該等替代符號值中的每一值都是藉由參考表或向量，而連結至基礎不變的 m 維事實資料表。然而，該矩陣呈現之維度 n 可等於該維度 m 。

基礎機制特別包含以下步驟：將屬於一明確資料索引鍵之所有資料值欄位的內容轉換成索引值 (充當索引值的記錄數字除外)；以及基於該等索引值，執行用於產生最終矩陣呈現之所有排序步驟，意即，整個排序機制或整個排序引擎僅使用該等索引值。

應提到，在例示性的二維狀況下，該 n 維矩陣呈現之 "基數" 意指該二維矩陣之行與列的數目。在更多維度 (意即 $n \geq 2$) 之狀況下，該基數相應地定義必要的呈現資源。

在一較佳實施例中，會將在一基礎矩陣呈現之單個單元格中所要呈現的資料記錄 (例如 2 維圖像矩陣之一列) 連續地排列成包含各自替代符號值之子序列。因此利用對應之該等子序列而產生樞紐分析呈現之該等單個單元格。因此，對一單元格之樞紐分析檢視而言，可使用資料記錄之整個序列的資料子序列。使用者所指定的樞紐分析呈現所需之該等單元格中的每一個特別定義為在該序列中的開始點，以藉由該開始點來存取待呈現之事實資料表所需的資料。

最終矩陣呈現藉由該資料序列而產生，藉此僅使用上述替代符號值來參考 n 維事實資料表。換言之，因此利用虛擬

2維矩陣來獲得該矩陣呈現。

在該較佳實施例中，藉由一包含較佳由整數組成之上述序列的線性向量而實現上述資料序列。由此，藉由該序列向量而重新排列每一維度(群組)之資料記錄，例如被包括在事實資料表中之資料記錄，以用於各自的使用者定義的樞紐分析組態。因為進而利用資料記錄索引數來完成對使用序列向量之真實資料記錄的存取，所以可即時(意即線上，意即無任何延遲或僅有非常短的延遲或使用者的回應時間)執行待被用於特殊樞紐分析呈現(意即待呈現於特殊樞紐分析檢視中之行與列)之單元格的計算。

根據另一實施例，首先將該基礎事實資料表轉換成一參考表，用以依據該參考表而產生該序列。在另一實施例中，會利用一計數向量，以決定資料記錄之該等維度中的兩個維度之間的轉換。

因此，根據本發明之機制不需要為了計算可能樞紐分析呈現的目的，(例如)藉由需要相當多資料處理與資料儲存資源的所有資料記錄預彙總，來預處理資料庫之資料記錄。

另外，歸因於中間序列化步驟，因而可擴大資料記錄之總量及因此可擴大資料庫大小，同時不會影擊可能樞紐分析呈現(檢視)，其中所需的處理時間單位隨基礎資料庫之大小而按線性比例變化。因此，第一維度映射至資料記錄之該等至少二維之至少第二維上。因此可互相映射資料庫之資料記錄的任意維度，因此使能夠藉由插入資料記錄之其他維或甚至其全部等級階層架構，來擴展現有的資料庫。

與以上討論之先前技術方法相比，根據本發明之機制不需要進行諸如大型資料集之間的比較作業來計算樞紐分析呈現。

另外，可在樞紐分析呈現之不同彙總階段中多次使用資料庫之特殊資料記錄維度，同時無任何執行時期的問題。因為僅對給定資料庫之其後的不同樞紐分析檢視使用向量運算，所以用於提供一使用者定義的樞紐分析檢視之資料記錄的先前已產生的序列保持有效，以用於其後的樞紐分析檢視，且因此其後的樞紐分析檢視總是由現有的樞紐分析檢視推斷而來，因此不破壞現有的樞紐分析檢視。

在本發明之另一態樣中，所選擇的2維樞紐分析呈現通常基於由行列數目確定的虛擬2維矩陣。提議的樞紐分析檢視處理機制因此為一給定的樞紐分析呈現提供行列數量，此外，還按上述子序列順序來提供行與列的交集點。因此，可僅藉由單元格之行與列索引來完成虛擬矩陣之該單元格的呈現。

因為藉由僅使用純參考和計數技術以及執行整數值之簡單線性(一維)向量運算，來產生或計算樞紐分析呈現，所以可利用資料庫記錄迅速產生一任意的使用者定義的樞紐分析呈現。

吾人強調根據本發明之樞紐分析呈現機制並不限於2維，其也可被應用於 n 維樞紐分析呈現。但是，在該 n 維狀況下，上述序列向量為 $(n-1)$ 維。

【實施方式】

下文中展示，對在通常狀況下由二維真實事實資料表所表示之給定的真實事實量而言，如何根據本發明來產生任意的樞紐分析表(交叉表)檢視。吾人強調該基礎樞紐分析表產生機制可用於為n維資料集產生m維矩陣呈現。另外，可用最少成本及消耗時間建構該機制，且僅需要最少計算資源，且因此使得產生高效能。

此處假設，真實事實量包括資料索引鍵(每一資料索引鍵定義一個索引鍵維度(key dimension))以及每一索引鍵維度中包含之資料值。檢視藉由以下各項提供事實量之樞紐分析：該等索引鍵維度之使用者所選擇的已排序排列、在該樞紐分析檢視之已呈現的維度中之索引鍵維度的排列及一可能的彙總層級，此在下文中更詳細地展示。

在圖1a與1b中展示，在當前狀況下由一真實事實資料表所表示之例示性事實量，及該真實事實資料表之例示性樞紐分析檢視。在該實例中，該真實事實資料表包含排列成索引鍵維度之銷售數字130、「顧客(Customer)」100、「物品(Article)」110及「月份(Month)」120。圖1b中展示之對應例示性樞紐分析檢視，按如下順序呈現圖1a所展示之真實事實資料表的(事實)資料值140：「顧客」/「物品」/「月份」150~170(圖中所示之左邊三行)，及銷售之總計值(右邊行)140。因此該等索引鍵維度，即「顧客」150與「物品」160在垂直方向呈現，其中索引鍵維度「月份」被彙總成日曆季度值「Q1」~「Q4」170，且呈現於在該樞紐分析檢視之水平方向。

產生該樞紐分析檢視之基礎機制的概念基礎為：將使用者所選擇之每一任意樞紐分析呈現歸納成為基礎事實量之資料值的連續排序排列(例如一序列向量)。圖2a至圖2c展示該機制之一較佳實施例。

在本實施例中，如圖2a所示，首先藉由一連續索引值200(左邊行)來擴展基礎真實事實資料表，在本實施例中，該索引值提供事實的自「1」至「20」的連續編號。

在圖2b所展示之所得樞紐分析檢視中，在每一單元格210中，該等事實之索引呈現在對應的單元格中必須求和的銷售值。

如上所述，基於一序列向量來產生圖2b所示之樞紐分析檢視。圖2c中描繪圖2b中之樞紐分析檢視的基礎序列向量，且由兩行260、270組成，左邊行260包含在本實施例中再次自「1」至「20」的連續數字，而右邊行270包含圖2a所描繪之上述索引值275，該等索引值係已排序的排列形式而得以按順序地建立圖2b之樞紐分析檢視。

為更好地理解序列向量，該序列向量之內容展示於相反方向，意即，自圖2b所描繪之樞紐分析檢視開始。因此，藉由在樞紐分析呈現之單元格210中對該等索引逐行且自左向右地求和，如在該樞紐分析檢視之前兩行中由箭頭250所展示，吾人將獲得圖2c所描繪之序列向量。

如上所述，僅藉由向量運算來產生樞紐分析檢視，而基礎向量僅包含整數值。因為以下描述的處理步驟均不使用實際的資料索引鍵值，如文字鍵值、本文或任何其他真實

事實值，所以在此獲得相當客觀的效能收益。在所有該等處理步驟中，僅使用上述整數(參考)值。僅當已完成該等整數處理步驟時，包含在所得序列向量中之參考值，藉由稍後所描述之指派或參考(查詢)表而轉換回對應的真實事實值。

在第一步驟中，圖3a中所描繪之真實事實資料表所示，包含在該事實資料表中之具有不同的索引鍵維度的所有資料(如顧客姓名"Miller"與"Brown")，被在該索引鍵維度中之唯一參考值300所取代，其中該等唯一參考值充當替代符號值使用。圖2a所示之例示性真實事實資料表的事實量在轉換為上述參考數字後，被展示於圖3a中。

在圖3b中展示用於轉換成參考數字之基礎機制。轉換係以數個指派(或參考)表350~370為基礎，其中會使用已知的查詢表機制，利用指派(或參考)表將具有不同的索引鍵維度之真實事實指派給上述參考值。如圖3b所示之索引鍵維度："顧客"350、"物品"360及"月份"370。如圖3c所展示之另一索引鍵維度"季度(Quarter)"380係充當一彙總等級階層架構使用，以使能夠進一步彙總索引鍵維度"月份"390。

對圖2a中所描繪之真實事實資料表的每一個索引鍵維度而言，第一整數向量410關於各個索引鍵維度與該等索引鍵維度之元素(參考值)的項目(occurrence)，來提供基礎事實行之事實排序次序。第二整數向量400提供一第一區塊元素(FBE)，該區塊元素為一個索引鍵維度元素(例如"Brown")之每一筆項目，指示在排序序列410中其對應區塊的第一元

素。現在參考圖4a至圖4c，在圖式中展示此。因此，在圖4a至圖4c中左邊所描述之資料表，在右邊呈現的排序向量中提供每一索引鍵維度的第一輸入項之參考索引。因此，例如圖4a所展示之索引鍵維度"顧客"，索引鍵"Brown"之所有事實數字420排列於行1至7中，索引鍵"Jones"之所有事實數字430排列於行8至14中，索引鍵"Miller"之所有事實數字440自行15排列至資料表的末端。

在該描述之末尾處更詳細地描述，用於獲得包含該排序次序之第一整數向量410與包含第一區塊元素索引之第二整數向量400所必需的方法。

值得注意的是，很容易利用"月份"之向量衍生而得索引鍵維度"季度"之對應表示，且因此在此處不再描述。

現在參考圖5及圖6a至圖6g，下文描述如何利用上述整數向量計算技術來排序該真實事實資料，以獲得所要的樞紐分析檢視。以提供一預排序步驟開始(該預排序步驟提供用於最終排序之事實的第一逐區塊排序)，逐步計算該排序(意即，為樞紐分析呈現的每個索引鍵維度計算一次)，以顯示最終的排序次序。藉由該等處理步驟，每個待被排序之索引鍵維度產生了新的群組，該等群組充當用於以下排序步驟，且用於在計算結束時提供樞紐分析呈現之單元格的內容的基礎。

在本實例中，藉由排序索引鍵維度"顧客"來顯示該等預排序區塊。基於相關向量，按圖5所展示的方式，將該等圖式所展示之命名為"映射(mapping)表"與"排序位置指標

(Sort position pointer)"表初始化。應強調，圖5所示之資料表僅為一快照集(snapshot)，該快照集描述按"顧客"對應"物品"對應"季度"的整個排序過程中之"顧客"對應"物品"的具體排序步驟。吾人將提到，遞迴地處理所有該等資料表，其中在每一排序步驟結束時，在各自的後續排序步驟中，在"映射"表與"排序位置指標"表中會重新使用包含在欄位"Fact# NewGrp"中之"結果排列"資料表的事實。然而，為了在第一排序步驟前初始化該等資料表，按以下方式初始化"映射"表與"排序位置指標"表之行。

在"顧客"之當前狀況下，意即利用該"真實事實"資料表之上述唯一參考值300，將排序序列之第一維度填入該"映射"表之"TmpGrp"行的每一欄位。另外，利用對應維度之上述FBE資料表的"Pos"行400來初始化該"排序位置指標"表之"NextPos"行的每一欄位。

在所有排序步驟期間，會將對應維度(意即在該對應排序序列步驟中之右邊維度)之"已排序序列"資料表的"事實(Fact)"行410填入"已排序序列"資料表之"Fact#"行。因此，在兩個上述初始化步驟之後清除"結果排列"內容，使得可在下一排序步驟中重新填充該等欄位。該資料表中之行(line)的數目等於包含在真實事實資料表中之事實的數目。以下進行的處理步驟由關於索引鍵維度"物品"之排序次序決定，該索引鍵的排序次序在該等圖式中命名為"已排序序列"。

在圖6a至圖6g所示之圖片序列中，更詳細地展示該排序

機制本身。該等圖片中的每一個圖片都展示單個排序步驟，該序列所示之排序步驟之整個序列因此僅描述整個排序程序之部分。該整個程序係以展示之四個資料表的內部連結為基礎。

首先，按自"1"至"20"之參數"Pseq"的次序，來處理索引鍵維度"顧客"之先前排序步驟而產生之索引鍵維度"物品"的資料表(名為"已排序序列")。具體地說，對包含在行"Fact#"中每一事實數字中而言，在該映射表中，確定一命名為"TmpGrp"之對應索引。

在"排序位置指標"表中使用圖6a所示之排序步驟中的所得索引(具體為所得的整數值"3")，以確定下一位置"NextPos"的值，意即，可藉由在一結果排列資料表中之目標索引來識別當前事實數字之值。另外，包含在該映射表之行"LDGrp"中的對應值與包含在"已排序序列"資料表中之對應值"DGrp"相比較。若該等兩個值不同，則該索引在"結果排列(Result Permutation)"資料表之行"開始(START)"中標記為"真(TRUE)"，因此定義一新的群組的開始。應提到，在圖6a中，在該程序之已展示狀態中尚未加入參數"NextPos"。

圖6b至圖6f展示另外5個步驟，且省略了步驟7至步驟19，圖6g展示命名為"結果排列"之最終所獲得的排列資料表(向量)。由此，在圖6g所示之"結果排列"資料表之"Fact#"行中，儲存"顧客"對應"物品"之排序步驟的結果。

在圖7中，再次展示圖6g所描繪之最終結果排列資料表。

其代表成相關排序序列排列之目標樞紐分析序列。在行 "NewGrp" 中，藉由自第一列開始將該群組索引增加 "1" 而為每一列指派一個值。若在行 "開始" 中，各個行標記為布林值 "真"，意即被標記為一新的群組的開始點，則完成增量之建立。該等群組係相關於由真實事實所提供之索引鍵維度 "顧客" 及 "物品" 之可能的不同組合。可在執行以上程序過程中容易地確定並管理對應於該等索引鍵維度中之該等群組的參考數字，而該等參考數字提供存取真實事實之最終呈現所需之各自基礎索引鍵本文。

進一步參考圖 8，其展示自先前排序步驟 "顧客" 對應 "物品" 之結果開始，初始化下一排序步驟 "顧客" 對應 "物品" 對應 "季度"。吾人提到包括一額外彙總步驟，其中將維度 "月份" 彙總成 "季度"。藉由一指示值項目的對應 "Pseq" 數字來初始化 "排序位置指標" 表之 "NextPos" 行，該指示值在本實施例中為 "結果排列" 資料表之 "開始" 行中之布林值 "真"。

自 "結果排列" 資料表之 "Fact#" 行開始，該行之每一欄位充當一進入點 (entry point) 使用，該進入點用於將圖 8 中所描繪之 "映射" 表之 "TmpGrp" 行中的 "一對應值填入" 結果排列 "資料表之 "NewGrp" 欄位中。

在圖 8 至圖 10 中，根據先前圖式來展示本實施例之最終排序步驟，因此顯示該最終樞紐分析呈現。因此，基於在行 "Fact#" 中提供之事實數字與在行 "NewGrp" 中提供之群組索引，來再次初始化 (如上所述) 另一對應之 "映射" 表及 "排序位置指標"，此使得產生索引鍵維度 "季度" 之下一排序序

列。在圖8中描述該初始化之結果資料表。在圖9a至圖9f中，展示最終排序序列之前5個步驟及最後排序步驟。該等圖片對應於相似的圖6b至圖9g，因此此處不再更詳細地描述。

整個排序序列，以計算圖10所描述之"樞紐分析序列"資料表而結束。以此方式，整個程序為已排序的索引鍵維度"顧客"對應"物品"對應"季度"的當前樞紐分析呈現揭示一次序序列向量。該次序序列向量儲存在"結果排列"資料表之"Fact#"行中。於行"NewGrp"中提供之基礎群組索引係關於基礎樞紐分析呈現或檢視的對應單元格。基於行"Fact#"中之各個輸入項可很容易地確定該等進一步對應事實參考。

上述給定實例僅由一個用於樞紐分析檢視之水平規格的索引鍵維度組成。應提到，該排序程序不限於任何數量的索引鍵維度，其中該等索引鍵維度用於樞紐分析規格(pivot specification)。

整個程序為每一樞紐分析維度確定一單獨排序序列。藉由將該排序程序應用至單獨產生之序列而產生整個序列，其中所述作業自第一序列之結果開始，且將下一序列作為該上述之"映射"表與"排序位置指標"表的輸入。

整個排序程序也適用於圖3a所示之"真實事實"資料表的子集。應提到，圖5所示之"映射"表之列的數目對應於"已排序序列"資料表中之"Fact#"行的最大可能值，意即，該"真實事實"資料表之事實列的數目。在僅有該等事實之一子集應用於該排序過程的狀況下(例如歸因於篩選準則)，可將

該"映射"表之大小縮減為已選擇之事實的大小。當"已排序序列"資料表之"Fact#"行的每一上述初始化，均使用對應子序列中之真實事實索引的相關索引時，可達成該項縮減。該排序過程之結果儲存在"結果排列"資料表之"Fact#"行中，而該"結果排列"資料表需要被轉換回真實事實索引。

如上所述，"第一區塊元素"(FBE)整數向量400與"排序次序"整數向量410係用於初始化每一排序步驟。計算關於圖3a所示之"真實事實"資料表之索引鍵維度行310的LBE索引向量。在第一步驟中，計算該計數向量，其中該計數向量之每一欄位包含該"真實事實#"資料表中之對應索引鍵維度之對應參考數字項目的數字。該計數向量中之每個值反映對應參考數字的區塊長度。處理自第一欄位至最後欄位求和該等值的計數向量，為每一步驟提供一對應群組之開始。該結果儲存在FBE索引向量400中。

利用該排序程序來計算"排序次序"整數向量410。在此狀況下，將相同值(例如"1")填入"已排序序列"資料表之"DGrp."行的每個欄位中。將"FSeq"行之對應值填入"Fact#"行之每個欄位。將圖3a所示之對應維度索引鍵的值填入"映射"表之"TmpGrp"行的每個欄位。將對應維度索引鍵之"FBE"整數向量400填入該"排序位置指標"表之"NextPos"行。一初始化即執行該排序程序，在"結果排列"資料表之"Fact#"行中提供該"排序次序"整數向量410。

應提到，當在當前樞紐分析規格內使用一階層層級(hierarchy level)(例如"季度")時會使用相同程序。當對應維

度索引鍵(例如"月份")之真實事實行轉換為該階層層級之參考數字時，可計算"FBE"與"排序次序"整數向量。

最後再次參考圖2b與圖7，其中更詳細地描述如何根據本發明來計算該等進入點。該最終樞紐分析呈現之基數(例如2維樞紐分析呈現中之行與列的數目)是排序程序之另一結果。對樞紐分析檢視之每個維度而言，對應樞紐分析檢視維度之基數視樞紐分析規格及事實而定，該等事實作為樞紐分析檢視之輸入。在給定的實例中，在"顧客"對應"物品"的排序步驟後，可衍生出垂直維度之基數。"結果排列"資料表之"NewGrp."行的最後欄位(參看圖6g)包含所得群組之數目。該數目相當於所得樞紐分析檢視之垂直維度的基數。因此，獨立地進行每一樞紐分析維度提供了對應樞紐分析維度所需的基數。

【圖式簡單說明】

前文中藉由參考附圖來更詳細地展示本發明，附圖使本發明之其他特點與優點變加明顯。

在以下圖式中：

圖1a及圖b描述在先前技術中已知的一例示性真實事實資料表，及一對應樞紐分析檢視；

圖2a至圖2c展示用於藉由資料表而產生根據本發明之樞紐分析檢視之機制的基本原理；

圖3a至圖3c展示用於產生根據本發明之樞紐分析檢視之機制的其他細節；

圖4a至圖4c展示用於產生根據本發明之序列向量的典型

向量運算；

圖 5 至 圖 6g 展示多個總覽資料表圖解(overview table diagram)，其用於展示如何儲存真實事實資料，以獲得根據本發明之所要的樞紐分析檢視。

圖 7 描繪根據本發明之最終結果排列資料表；以及

圖 8 至 圖 10 描述根據本發明之最終排序步驟。

【圖式代表符號說明】

- 100 顧客
- 110 物品
- 120 月份
- 130 銷售數字
- 140 資料值
- 150 顧客
- 160 物品
- 170 月份
- 200 索引值
- 210 單元格
- 250 箭頭
- 260 左邊行
- 270 右邊行
- 275 索引值
- 300 唯一參考值
- 310 索引鍵維度行
- 350 顧客

360 物 品

370 月 份

380 季 度

390 月 份

400 整 數 向 量

410 整 數 向 量

420 事 實 數 字

430 事 實 數 字

440 事 實 數 字

伍、中文發明摘要：

本發明揭示一種用於產生被包含在一資料庫中之資料記錄的使用者定義之樞紐分析檢視(pivot view)的方法及系統，在該資料庫中(如圖2a所示)由連續的索引值(index value)(200)來擴展一基礎真實事實(real fact)資料表，該索引值自「1」至「x」(在本實例中 $x=20$)提供該等事實之連續編號。在圖2b所展示之所得樞紐分析檢視中，在每一單元格(210)中，該等事實之索引被表示為：必須在對應的單元格中求和的銷售值。藉由一序列向量而產生該樞紐分析檢視。圖2c中描繪圖2b中之樞紐分析檢視的基礎序列向量且由兩行(260, 270)組成，左邊行(260)包含又是自「1」至「x」之連續數字，右邊行(270)包含圖2a中描繪之已提到的索引值(275)，該等索引值係已排序的排列形式而得以按順序地建立圖2b之樞紐分析檢視。

陸、英文發明摘要：

Disclosed are a method and system for generating user-defined pivot views of data records contained in a database where, as depicted in Fig. 2a, an underlying real facts table at first is extended by continuous index values (200) which provides a continuous numbering of the facts from ,1' to ,x' ($x = 20$ in the present example). In the resulting pivot view shown in Fig. 2b, in each cell (210) the indices of those facts are presented which sales value has to be summed-up in the corresponding cell. The pivot view is generated by means of a sequence vector. The underlying sequence vector for the pivot view in Fig. 2b is depicted in Fig. 2c and consists of two columns (260, 270), the left column (260) containing continuous numbers from again ,1' to ,x' and the right column (270) containing the mentioned index values (275) depicted in Fig. 2a in an ordered arrangement that enables sequentially building-up the pivot view of Fig. 2b. (Fig. 2)

柒、指定代表圖：

(一)本案指定代表圖為：第（ 2 ）圖。

(二)本代表圖之元件代表符號簡單說明：

200 索引值

210 單元格

250 箭頭

260 左邊行

270 右邊行

275 索引值

捌、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

（無）

拾壹、圖式：

圖 1A

Real Facts			
Customer	Article	Month	Sales
Miller	a4711	Jul	30
Brown	a4711	Jan	60
Miller	a4711	Jan	30
Miller	a4711	May	70
Miller	a0815	Dec	30
Jones	a4711	Jan	50
Jones	a0815	Aug	50
Brown	a0815	Oct	60
Jones	a4711	Nov	20
Miller	a4711	Mar	40
Brown	a4711	Sep	10
Jones	a0815	Feb	60
Brown	a0815	Apr	30
Brown	a4711	Dec	30
Brown	a0815	Feb	50
Jones	a0815	Sep	30
Jones	a0815	Mar	50
Brown	a0815	Nov	20
Miller	a0815	Apr	70
Jones	a4711	Jun	10

圖 1B

Pivot View			
Customer	Article	Sales	
Brown	a0815	50	30
Jones	a0815	60	10
Jones	a4711	110	80
Miller	a0815	50	10
Miller	a4711	70	30

Pivot Sequence			
PSeq	Fact#	Stat	NewGrp
1	15	True	1
2	13	True	2
3	8	True	3
4	18		3
5	2	True	4
6	11	True	5
7	14	True	6
8	12	True	7
9	17		7
10	7	True	8
11	16		8
12	6	True	9
13	20	True	10
14	9	True	11
15	19	True	12
16	5	True	13
17	3	True	14
18	10		14
19	4	True	15
20	1	True	16

圖 10

200

Real Facts				
Fact	Customer	Article	Month	Sales
1	Miller	a4711	Jul	30
2	Brown	a4711	Jan	60
3	Miller	a4711	Jan	30
4	Miller	a4711	May	70
5	Miller	a0815	Dec	30
6	Jones	a4711	Jan	50
7	Jones	a0815	Aug	50
8	Brown	a0815	Oct	60
9	Jones	a4711	Nov	20
10	Miller	a4711	Mar	40
11	Brown	a4711	Sep	10
12	Jones	a0815	Feb	60
13	Brown	a0815	Apr	30
14	Brown	a4711	Dec	30
15	Brown	a0815	Feb	50
16	Jones	a0815	Sep	30
17	Jones	a0815	Mar	50
18	Brown	a0815	Nov	20
19	Miller	a0815	Apr	70
20	Jones	a4711	Jun	10

圖 2A

Pivot view				
	Customer	Article	Sales	
			01	02
			03	04
	Brown	a0815	15	13
		a4711	2	11
	Jones	a0815	12.47	7.16
		a4711	6	20
		a0815		19
	Miller	a4711	3.10	4
				1
				5

250

210

圖 2B

Pivot Sequence	
PSeq	Fact#
1	15
2	13
3	8
4	18
5	2
6	11
7	14
8	12
9	17
10	7
11	16
12	6
13	20
14	9
15	19
16	5
17	3
18	10
19	4
20	1

275

260

270

圖 2C

310

Real Facts				
Fact	Customer	Article	Month	Sales
1	3	2	7	3
2	1	2	1	6
3	3	2	1	3
4	3	2	5	7
5	3	1	12	3
6	2	2	1	5
7	2	1	8	5
8	1	1	10	6
9	2	2	11	2
10	3	2	3	4
11	1	2	9	1
12	2	1	2	6
13	1	1	4	3
14	1	2	12	3
15	1	1	2	5
16	2	1	9	3
17	2	1	3	5
18	1	1	11	2
19	3	1	4	7
20	2	2	6	1

300

圖 3A

350

Dim	Customer
Ref	CustomerKey
1	Brown
2	Jones
3	Miller

360

Dim	Article
Ref	ArticleKey
1	a0815
2	a4711

370

Dim	Month
Ref	MonthKey
1	Jan
2	Feb
3	Mar
4	Apr
5	May
6	Jun
7	Jul
8	Aug
9	Sep
10	Oct
11	Nov
12	Dec

圖 3B

390

Dim	Month	QuarterKey
Ref	MonthKey	QuarterKey
1	Jan	Q1
2	Feb	Q1
3	Mar	Q1
4	Apr	Q2
5	May	Q2
6	Jun	Q2
7	Jul	Q3
8	Aug	Q3
9	Sep	Q3
10	Oct	Q4
11	Nov	Q4
12	Dec	Q4

380

Dim	Quarter
Ref	QuarterKey
1	Q1
2	Q2
3	Q3
4	Q4

圖 3C

圖 4A

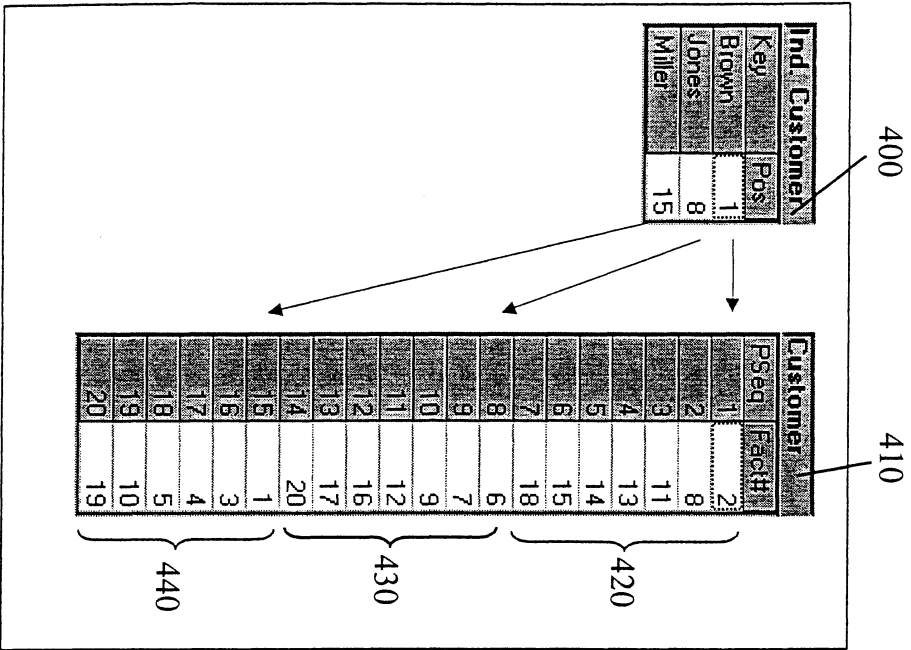


圖 4B

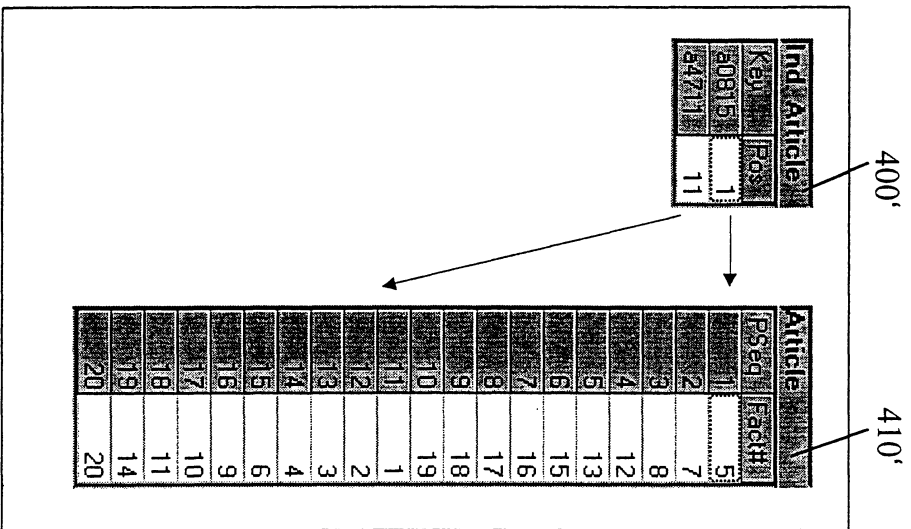
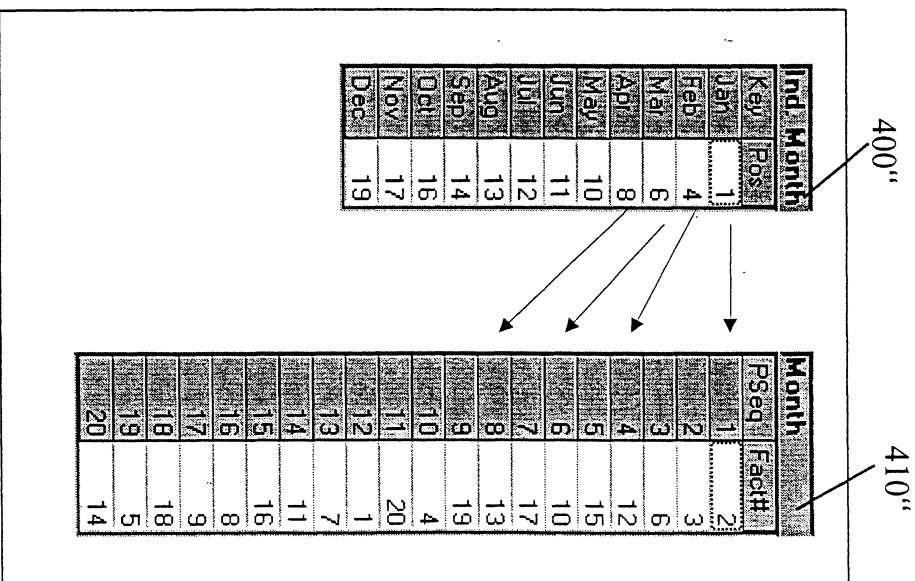


圖 4C



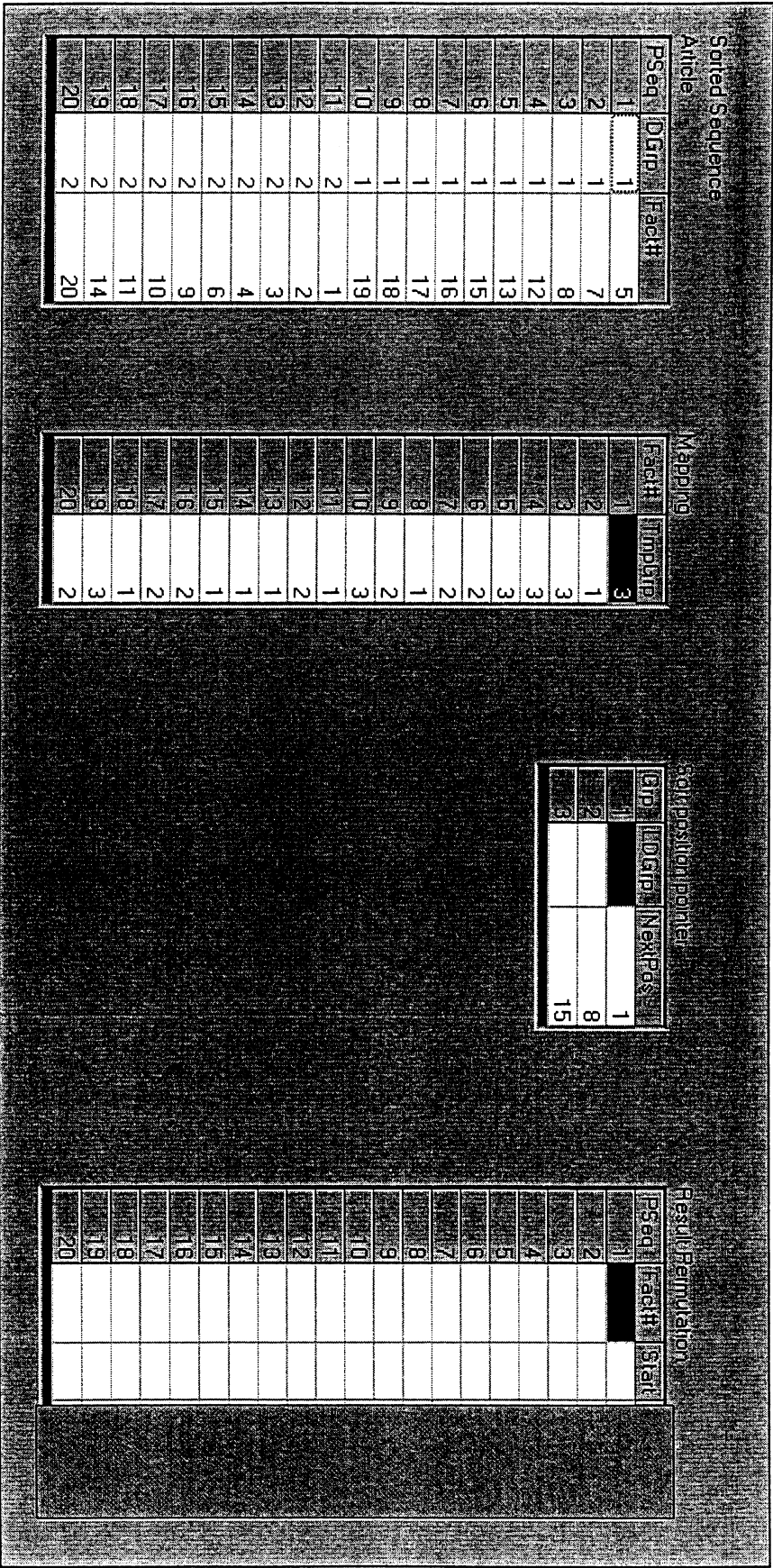


圖 5

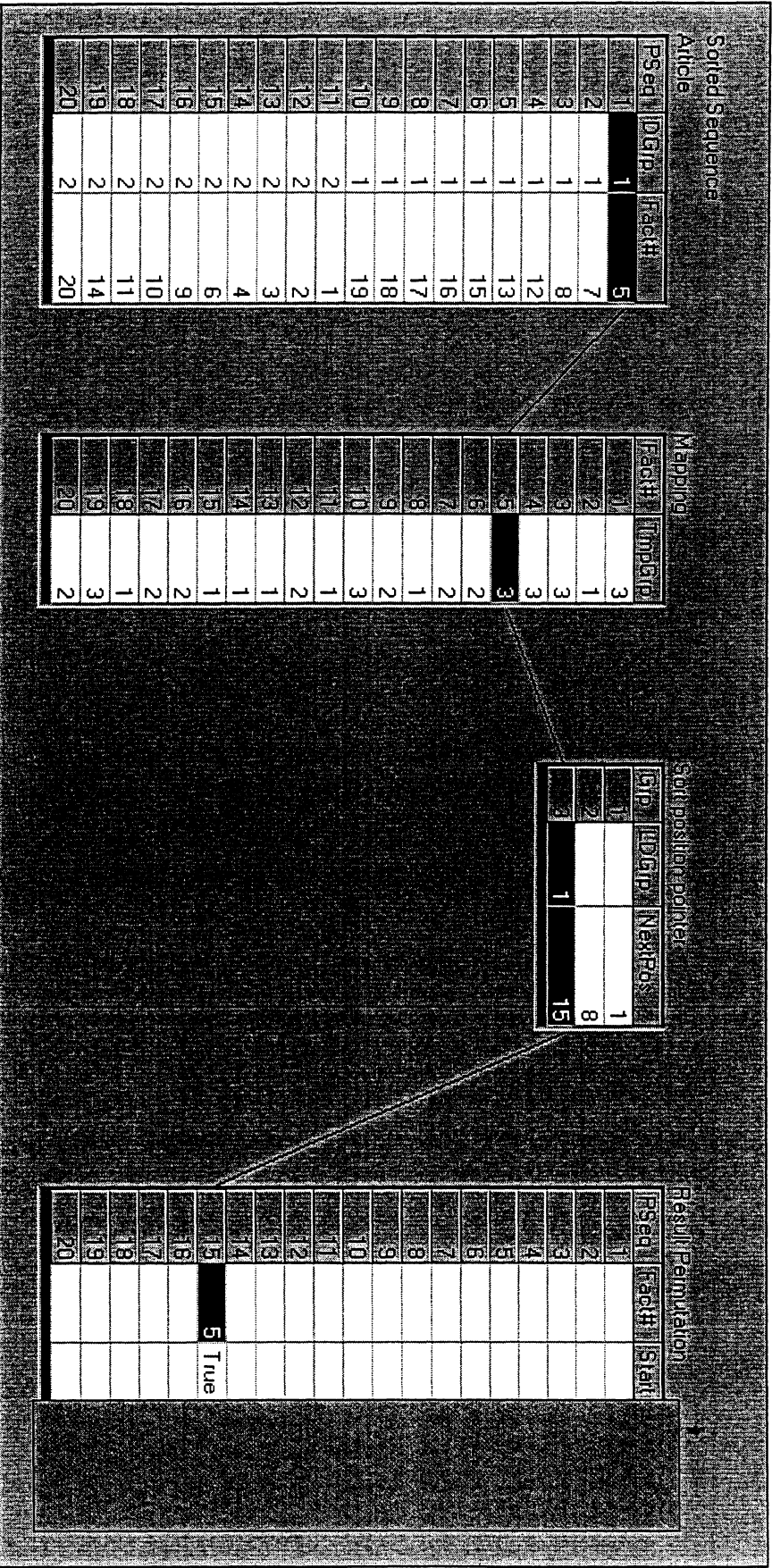


圖 6A

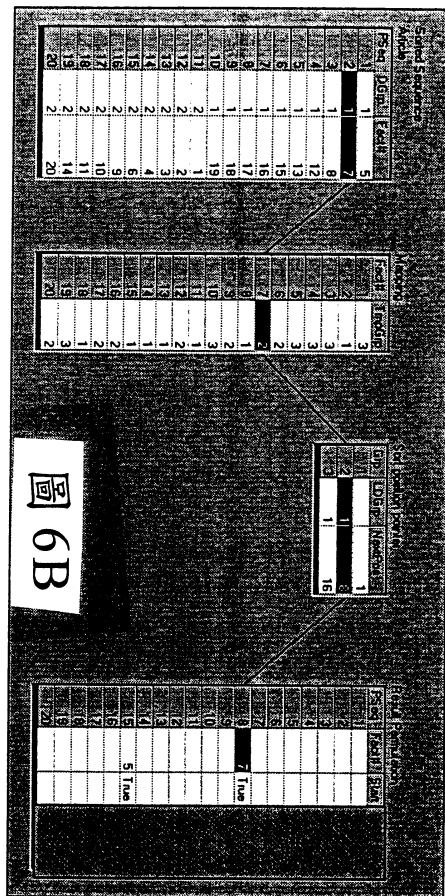


圖 6B

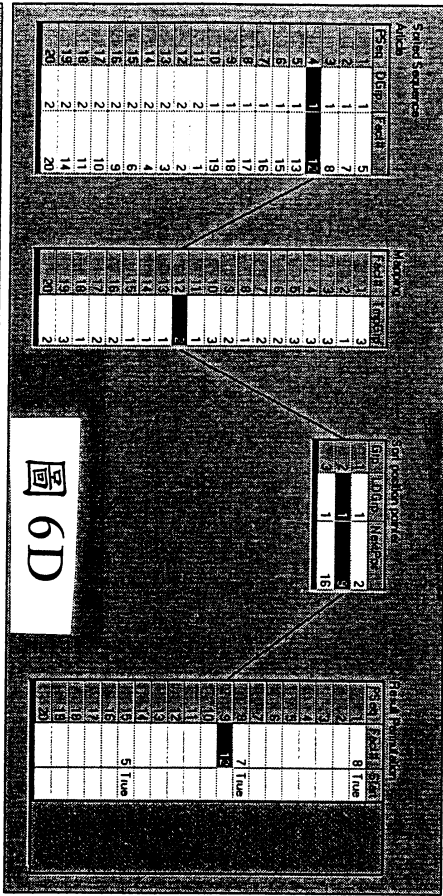


圖 6D

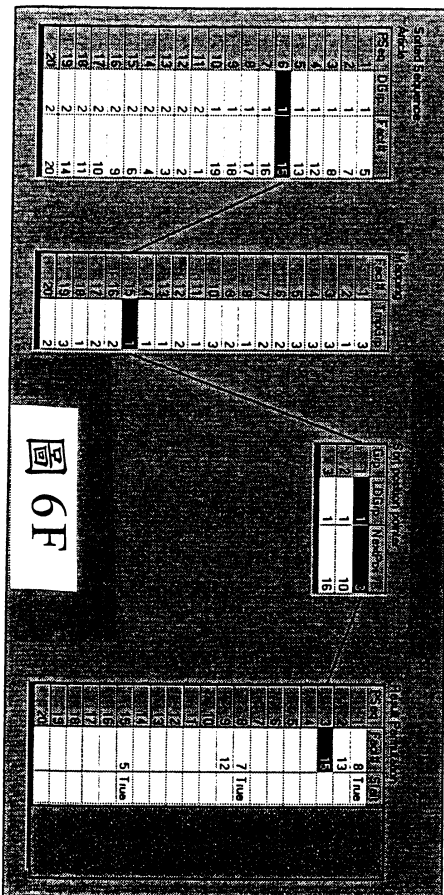


圖 6F

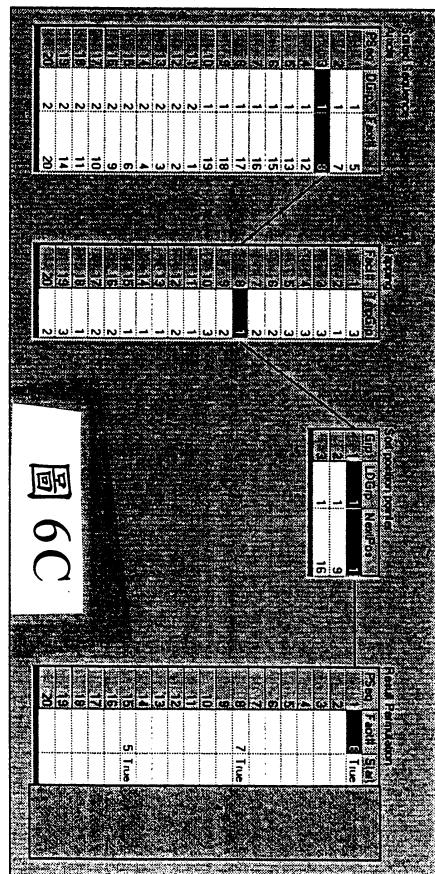


圖 6C

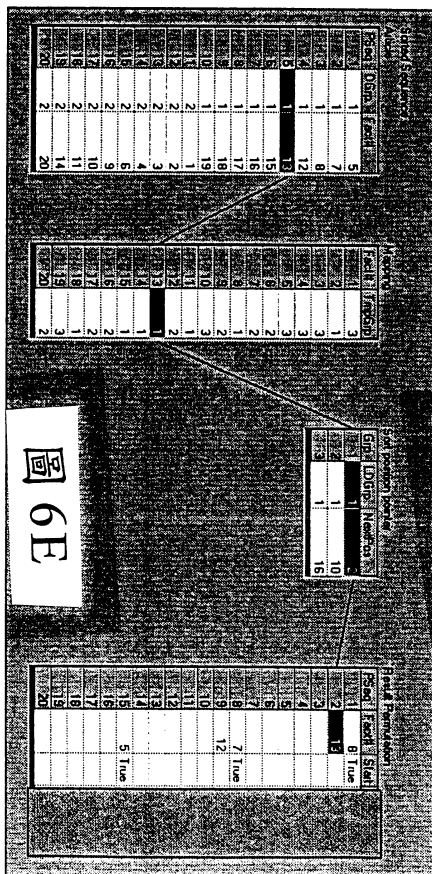


圖 6E

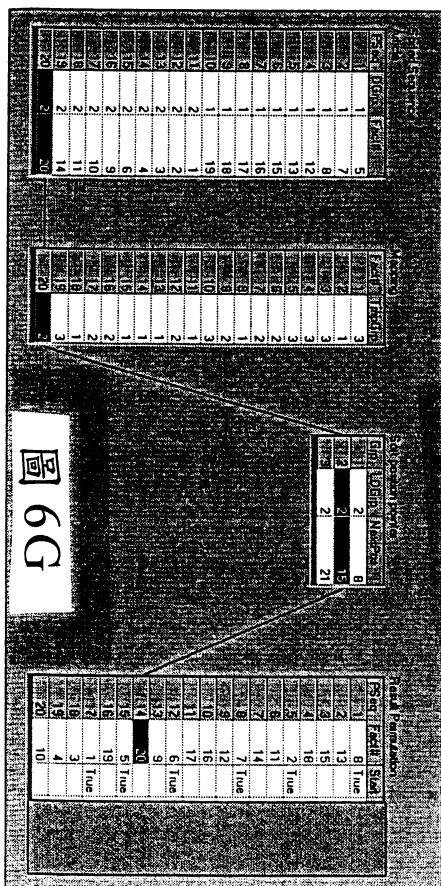


圖 6G

Result Permutation				
PSeq	Facit	Stat	Navgrp	
1	8	True	1	
2	13		1	
3	15		1	
4	18		1	
5	2	True	2	
6	11		2	
7	14		2	
8	7	True	3	
9	12		3	
10	16		3	
11	17		3	
12	6	True	4	
13	9		4	
14	20		4	
15	5	True	5	
16	19		5	
17	1	True	6	
18	3		6	
19	4		6	
20	10		6	

圖 7

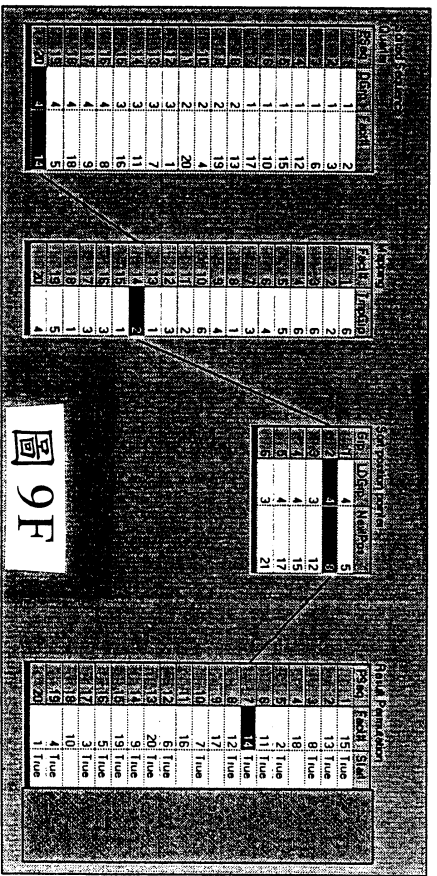
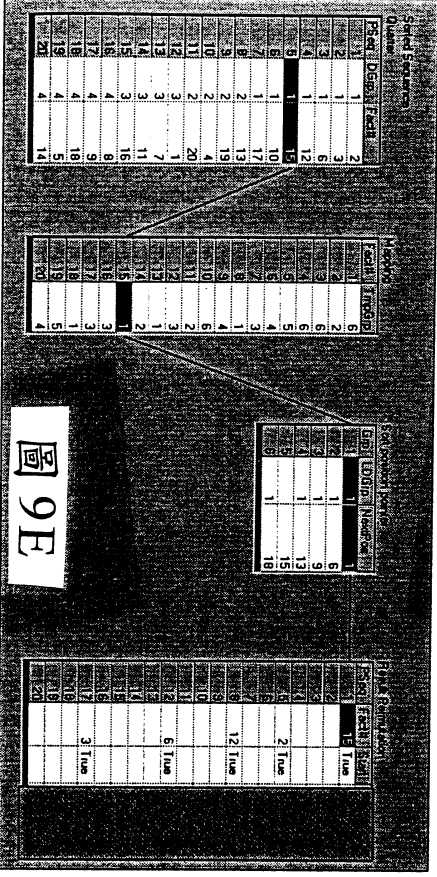
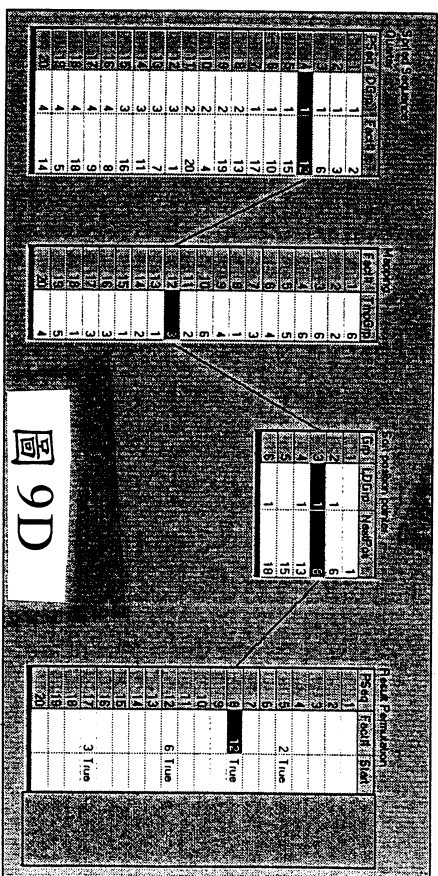
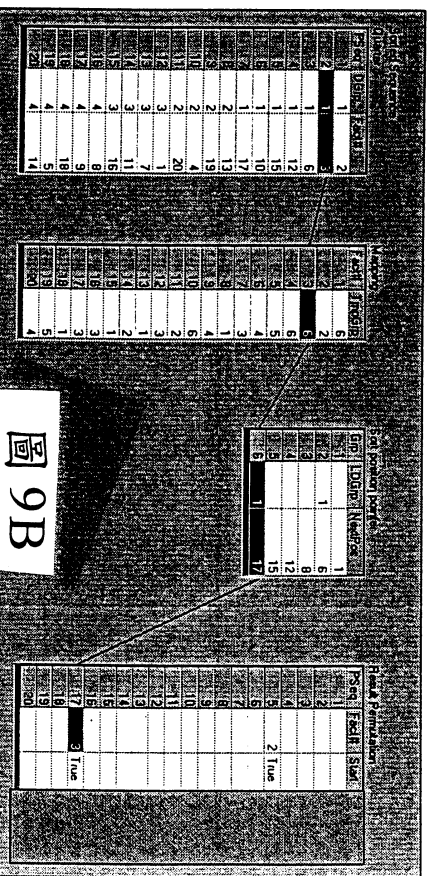
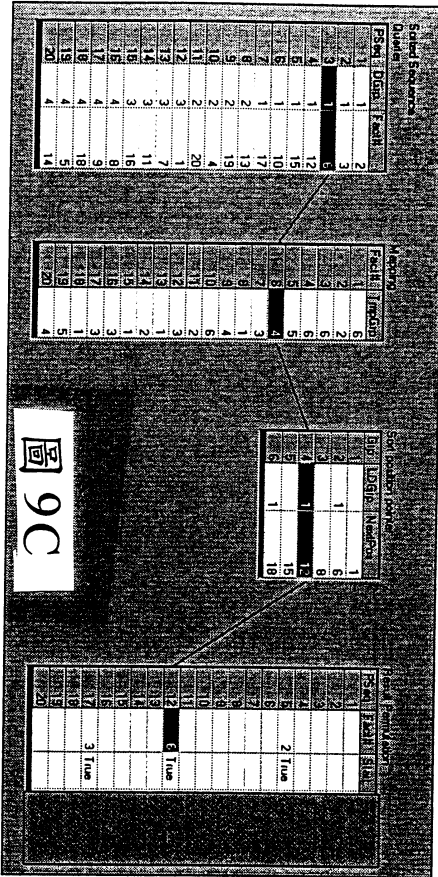
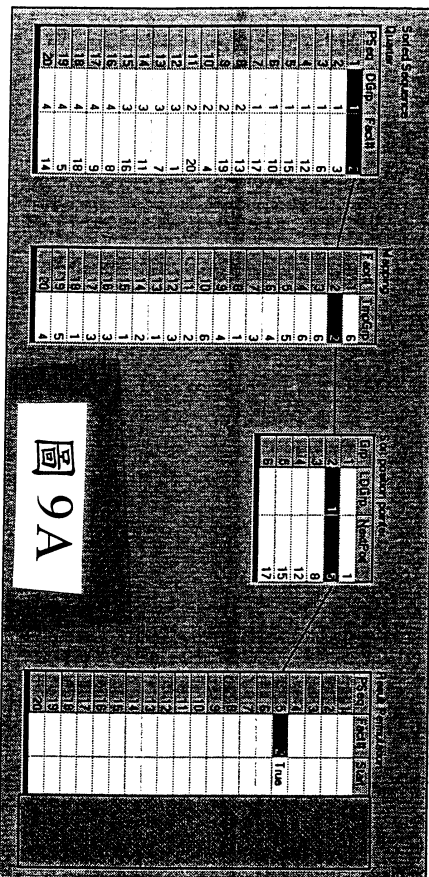
Sorted Sequence				
PSeq	Orig	Facit		
1	1	2		
2	1	3		
3	1	6		
4	1	12		
5	1	15		
6	1	10		
7	1	17		
8	2	13		
9	2	19		
10	2	4		
11	2	20		
12	3	1		
13	3	7		
14	3	11		
15	3	16		
16	4	8		
17	4	9		
18	4	18		
19	4	5		
20	4	14		

Mapping				
PSeq	Facit	Invseq		
1	6			
2	2			
3	6			
4	6			
5	5			
6	4			
7	3			
8	1			
9	4			
10	6			
11	2			
12	3			
13	1			
14	2			
15	1			
16	3			
17	3			
18	1			
19	5			
20	4			

Subposition point				
Orig	LDGIP	Navseq		
1		1		
2		5		
3		8		
4		12		
5		15		
6		17		

Result Permutation				
PSeq	Facit	Stat		
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

圖 8





拾、申請專利範圍：

1. 一種電腦所建構的方法，其用於產生一 m 維資料庫之至少一部分之一任意 n 維矩陣呈現，該 m 維資料庫由具有至少一索引鍵維度與至少一對應資料值欄位的多筆真實資料記錄組成，

該方法包含以下步驟：

根據該 n 維矩陣呈現，來計算該等資料值欄位之多個輸入項的一排序序列；

對該 n 維矩陣呈現之每一維度而言，將對應進入點資訊計算成該排序序列；

基於該已計算的排序序列與該已計算的進入點資訊，來計算該 n 維矩陣呈現之基數；

基於該基數而產生該任意 n 維矩陣呈現，其中基於該排序序列與該進入點資訊來處理該等資料值欄位。

2. 如申請專利範圍第1項之方法，其中將所有資料記錄之該等索引鍵維度轉換為索引值，該等索引值參考該等索引鍵維度之對應鍵項目的屬性，且其中將多個已排序之索引值重新轉換回該真實資料，以產生該任意矩陣呈現。
3. 如申請專利範圍第2項之方法，其中由一向量來表示一個索引鍵維度之該等索引值。
4. 如申請專利範圍第2項或第3項之方法，其中由整數來表示該等索引值。
5. 如申請專利範圍第2項或第3項之方法，包含以下特殊步

驟：

為該資料庫之每一維度，產生一專門由該等整數值組成的參考表，其中該參考表包括每一資料記錄之多個記錄識別項，由此保持該資料庫之維度結構；

基於該參考表，藉由基於該任意矩陣呈現之維度結構的純量向量運算(scalar vector operation)，來產生一序列向量；

基於該參考表產生一提供資料記錄之該至少一維度中包含的資料記錄量的計數向量，因此定義資料之不同的該至少一維度間的轉移；

呈現該資料庫之該任意矩陣呈現，其中利用該等資料識別項，使用包含在該序列向量與該計數向量中之該真實資料，來使多個資料記錄之該至少一維度形成群組。

6. 如申請專利範圍第5項之方法，其中該序列向量僅包含整數。
7. 如申請專利範圍第2項或第3項之方法，其中會基於一第一區塊元素(FBE)索引及/或一最後區塊元素(LBE)索引來計算該進入點資訊之。
8. 如申請專利範圍第2項或第3項之方法，其中將真實資料記錄之該至少一維度的第一維度，映射至真實資料記錄之該至少一維度的至少一第二維度。
9. 如申請專利範圍第2項或第3項之方法，其中基於一真實事實資料表而產生該參考表，首先會藉由提供該等真實事實之一連續編號的連續索引值來擴展該真實事實資

料表。

10. 如申請專利範圍第2項或第3項之方法，其中該序列向量由兩行組成，左邊行包含該等連續整數，且該右邊行(270)包含已排序排列形式的該等索引值，而得以按順序建立該任意矩陣呈現。
11. 如申請專利範圍第2項或第3項之方法，其中僅執行基於整數之向量計算，以產生該任意矩陣呈現。
12. 如申請專利範圍第2項或第3項之方法，其中執行一預排序步驟，該預排序步驟提供用於稍後排序步驟之該等真實事實的第一逐區塊排序，其中對該等區塊中之該樞紐分析呈現的每個索引鍵維度，逐步計算該排序，以揭示該任意排序次序。
13. 如申請專利範圍第2項或第3項之方法，其中在一排序位置指標表中使用由該等排序步驟所得之一索引值，以確定一結果排列資料表中之目標索引值，可藉由該目標索引值來識別一當前真實事實值的值。
14. 如申請專利範圍第13項之方法，其中該結果排列資料表表示一相關排序序列形式的目標矩陣序列。
15. 一種儲存在一電腦可使用媒體上之電腦程式產品，該電腦程式產品包含多個軟體程式碼構件，當在一電腦上執行時該程式時，該等軟體程式碼構件促使該電腦執行如申請專利範圍第1項至第14項中任一項之方法。
16. 一種線上分析處理(OLAP)系統，其用於藉由一資料庫之至少部分的多個任意 n 維矩陣呈現，來分析包含在該

資料庫中之真實事實資料，其中該資料庫包含多個資料記錄的至少一維度，每個資料記錄都包含至少一資料索引鍵及一對應資料值欄位，該OLAP系統之特徵為如申請專利範圍第1項之該方法。