

## 混合式資訊推薦於音樂 CD 資訊之研究

### Hybrid Information Recommendation System for Music CD

李來錫<sup>1</sup> 黃文益<sup>2</sup>

(Received: Oct. 26, 2007 ; First Revision: Dec. 15, 2007 ; Accepted: Mar. 20, 2008)

---

#### 摘要

音樂CD種類繁多，在挑選上常須耗費相當多的時間。混合式推薦系統結合協同式與內容式推薦系統的特性，可以協助使用者從大量的產品資訊中找出想要的產品。本研究以混合式資訊推薦方法進行音樂CD資訊的推薦工作，其中協同式推薦是藉由過去銷售記錄，挑選出使用者購買相同偏好的CD，再以內容式推薦依使用者偏好做二次篩選協同式推薦之結果。研究結果顯示，混合式資訊推薦可提供更精確的產品篩選程序，並獲得良好的推薦效果。

**關鍵詞：**混合式推薦系統、協同式推薦系統、內容式推薦系統

#### Abstract

Selecting music CDs is an interesting work seeking to meet our changing preferences on kinds of music. However, it takes much time to select favorite music CDs from lots of CDs. This study aims to develop a hybrid information recommendation system which combines the collaborative filtering technique and the content-based filtering technique. The collaborative filtering technique can filter music CDs based on the preferences of near neighbors. And the content-based filtering will filter music CDs at the second stage based uses own preferences. The result shows the hybrid information system can be used to assist users for selecting music CDs.

**Keywords :** Hybrid Recommendation System, Collaborative Recommendation System, Content-based Recommendation System

---

<sup>1</sup>國立屏東商業技術學院資訊管理系助理教授

<sup>2</sup>國立屏東商業技術學院資訊管理系研究生

## 1. 緒論

網際網路已成為現代人接收資訊的主要媒介之一，然而過多的資訊量卻產生資訊超載 (information overload) 的問題，因此為了克服產品資訊超載的問題，許多網站業者為求提升使用者的網路資訊服務，便不斷改善資訊推薦系統的功能，以期滿足使用者需求。

目前資訊推薦系統技術主要歸納成二種：內容式推薦系統 (content-based recommendation system) 及協同式推薦系統 (collaborative recommendation system) (Wang and Shao, 2004)。內容式推薦系統是利用使用者過去的使用習慣及偏好，進而篩選出適合的資訊給使用者，即藉由分析使用者的特性來建立使用者輪廓 (profile)，然後計算出特定的產品內容與使用者輪廓 (profile) 之相似度，推薦高相似度的產品給使用者參考。而協同式推薦系統是透過群體的觀點來產生推薦項目給目標使用者，故強調在具有相同偏好的社群中，藉由過去使用者之間的歷史記錄，計算出各使用者之間偏好行為的相似度，以找出與自己偏好最相近的鄰居 (nearest neighbors)，並透過這些最近鄰居所組成的鄰近群組之意見，建議資訊給目標使用者做為參考。

現今市面上音樂 CD 的種類繁多且複雜，想要在眾多音樂 CD 中挑選合適的產品常得耗費相當多的時間，因此可利用資訊推薦系統來降低挑選的時間。內容式推薦系統和協同式推薦系統等推薦技術均有各自的限制與缺點，因此後續有許多學者提出混合式推薦模型，以互相補助其缺點的部份。Kim and Kim (2001) 曾指出混合式的推薦模型，不僅可保有以上二種個別推薦模式之優點，且其推薦之品質亦較內容式和協同式之個別推薦模式來得較佳。因此本研究嘗試利用混合式推薦機制應用在音樂 CD 推薦上，希望藉由此機制篩選出使用者喜愛的音樂 CD，以減少使用者搜尋的時間，並輔助其購買決策的制定。

## 2. 文獻探討

本研究主要的目的是運用混合式資訊推薦方法，以協助使用者找尋其想要或合適的產品。以下針對內容式推薦系統、協同式推薦系統與混合式推薦系統三個部分，依序作相關文獻探討。

### 2.1 內容式推薦系統

內容式推薦的觀念是從資訊檢索 (information retrieval) 領域衍生而來，主要是比對使用者偏好與項目內容之間的相關性，以提供使用者建議的一種推薦方法。此方法必須先建立使用者輪廓，而使用者輪廓是由一些關鍵詞或者是相關文件所組成。內容式推薦系統會將每一份新加入的文件與其他文件比對，並推薦給使用者相關之內容，如 Verity 的 Topic 即是採用此類技術 (卜小蝶, 1998)。

許多學者將內容式推薦應用於實務上，Lang (1995) 曾提出一個網路新聞推薦系統 NewsWeeder，該系統將使用者瀏覽過的文章由使用者進行評筆，再將評比資料整理分

析且建立使用者輪廓資料檔；Armstrong 等人（1995）也提出網際網路資訊搜尋系統 WebWatcher，該系統可透過使用者輸入有興趣之主題，以協助在相關特定網站內找到所需要的資訊。Krulwich and Burkey（1997）也發展出類似的系統 InfoFinder，可利用訊息資料集（message sets）或者是其他線上文件，來預測使用者偏好類別。InfoFinder 的特點在於利用啟發式（heuristic）搜尋技巧來取得有意義的片語，此方法的優點是不需要很多文件樣本，便可以正確找出使用者的興趣；Lawrence 等人（2001）則提出系統 SmartPad，以每位使用者對產品類別的偏好與產品之間的相似度，來提供個人化之產品推薦清單。

雖然內容式雖然內容式推薦系統已經被發展成數套系統，而且可實際運作，但是此類系統仍存在三個主要的缺點（Shardanand and Maes, 1995；Balabanovic and Shoham, 1997）：

1. 要能以文字表達出來才能被內容式推薦系統分析，例如藝術、電影、旅遊等，無法以文字表達之產品即無法利用此方法。
2. 因為內容式推薦系統是依使用者過去之偏好為依據，所以，僅能推薦與過去相類似的產品，而無法推薦未曾接觸過但卻有可能非常喜愛或有興趣之產品。
3. 無法篩選品質（quality）、風格（style）與觀點（point of view）。例如兩篇主題相同，可是內容品質、風格以及觀點各不相同的文章，內容式推薦系統就不能篩選這兩篇文章。

基於以上所敘述的內容式推薦缺點，故近年來有越來越多的學者，提出協同式推薦系統，以解決內容式推薦系統無法解決的問題。

## 2.2 協同式推薦系統

協同式推薦可記錄使用者偏好，並分類成不同群體（groups），而且利用群體合作方式來推薦相關的資訊。相關的系統有 Goldberg 等人（1992）提出電子郵件分類系統 TAPESTRY，後來延伸到所有電子文件；Shardanand and Maes（1995）也提出類似的系統 Ringo，可利用使用者之偏好對演唱者以相關等級作評比，而且以這些評比做為使用者輪廓資料，最後 Ringo 再藉由輪廓資料決定偏好相似的使用者，並以此做為推薦依據；Konstan 等人（1997）發展出系統 GroupLens，可用來過濾 Usenet 的新聞，透過將使用者看過的新聞與文章進行評分，再利用 Pearson 相關係數計算使用者之間的相似度，以找尋合適的鄰居群組，並且主動推薦使用者感興趣的新聞或文章；同時，Terveen 等人（1997）開發出系統 PHOAKS，藉由隱性回饋來推論使用者偏好，即利用 Netnews 文章中較好的網站 URL 推薦給其他使用者，以及分析張貼之資訊，篩選出資訊中推薦的網站 URL 與統計每一個網站 URL 有多少人推薦，將相關網站 URL 推薦給需要的使用者。協同式推薦會優先考慮使用者之偏好，藉由相同群體的評價來進行推薦，再利用同群體的偏好來預測未知使用者的偏好，即可透過同群體曾經購買或觀看過的資訊之評分高低來進行推薦。

協同式推薦方法是將相同喜好的使用者分成同一群體，再利用相同群體的評價來推測其他使用者的偏好。許多學者提出相關推薦方法來改善使用者分群與資訊篩選，如

Chai and Vercoe (2000) 提出以使用者模型為基礎之音樂推薦代理人，使用者模型是由使用者自行定義，記錄使用者習慣、喜好以及其他使用者的特徵，利用使用者模型從擁有的音樂中選擇適合的音樂來推薦。Chen and Chen (2001) 則是依使用者偏好與音樂特徵分群來推薦，所選擇的音樂特徵包括音高 (pitch)、速度 (tempo degree) 以及音量 (loudness) 等。Hayes and Cunningham (2004) 建立一個自動化的協同式篩選系統 (automated collaborative filtering)，運用線上即時 (real-time) 的方式擷取使用者喜愛的音樂類型進行推薦。綜合上述協同式推薦的方法，各系統的做法皆先記錄過去使用者的偏好與習慣來進行分群，再以相同群體的評價來推測其他使用者的偏好。Roh 等人 (2003) 以及 Chao and Kim (2004) 曾提出協同式推薦是目前使用最成功之推薦技術，但仍然存在下列三個主要的缺點：

### 1. 稀疏性問題

稀疏性問題 (sparsity problem)，即使用者評比的產品數量太多，會造成評比資料矩陣過於稀疏，因此資訊推薦系統在篩選時便無法找到合適的產品推薦給使用者。若要提高評比資料矩陣密度的方法就是讓使用者評比更多的產品，但是會增加使用者的負擔。

### 2. 冷啟始問題

冷啟始問題 (cold-start problem)，因為系統開始啟用期間，新進使用者缺乏足夠的評比資料來做為分群的依據，所以無法將使用者歸類到某一個群組，而且系統也無法推薦資訊給使用者參考。必須蒐集足夠的使用者評比後，才能將使用者分群，給予合適的推薦。

### 3. 擴充性問題

擴充性問題 (scalability problem)，則是透過最相近鄰居演算法來計算使用者之間的相似度，其演算法的時間複雜度，是與使用者歷史交易資料，以及評比項目數量的多寡成等比平方，因此當使用者與推薦項目逐日增多的時候，系統擴充性相對的就會出現困難。

## 2.3 混合式推薦系統

雖然協同式推薦技術的發展，一直被認為是最成功的推薦系統核心技術，但隨著研究推展以及實際應用上遭遇到各種問題，因此有學者將提出以混合式推薦設計做為推薦系統的核心，而所謂混合式推薦是結合兩種或兩種以上不同的資訊篩選技術所構成，一般常見的混合方式即為結合內容式推薦與協同式推薦來實作。然而也有結合不同演算機制所衍生的混合式推薦，皆是為了試圖解決預測不準確與計算效能的瓶頸問題。

混合式推薦系統主要分為兩大類，第一類是將兩種推薦方法的結果整合在一起，如 Wasfi 等人 (1999) 曾提出系統 ProfBuilder，分別以協同式與內容式推薦產生兩組推薦清單，協同式推薦是比對使用者與過去相同群體瀏覽之網站路徑，來推薦相關的網站，而內容式推薦是依使用者瀏覽網站的內容與偏好，來推薦相似性高的網站。另外，Claypool 等人 (1999) 也曾提出系統 Personal Tango，同樣將協同式與內容式推薦分別獨立出來，而相似度是由兩者個別的權重相乘，來進行推薦。第二類是將兩種推薦方法

整合在一起產生一組推薦結果，如 Balabanovic and Shoham (1997) 提出系統 Fab，它混合協同式與內容式推薦兩種不同的方法，為網站讀者推薦喜愛的文章，且記錄每一位讀者的喜好，以找出相同偏好，再以協同式推薦技術推薦文章給每一位讀者。

近年來，各種混合式的推薦技術持續被提出，主要的作法仍以先瞭解使用者習慣與偏好，再進行相關的資訊推薦。例如，Liu and Shih (2005) 曾結合消費者的偏好與其消費能力，以求取消費者最可能購買的物品；Choi 等人(2006)則重從產品本身的特殊價值，探討消費者的購買意願，進而進行資訊推薦；Cao and Li (2007) 與 Zhang and Jiao(2007) 皆認為要進行有效的資訊推薦，必須先瞭解所售產品的屬性，再進行有效分類後，才能進行可以滿足消費者需求的資訊推薦；Li 等人(2007)也曾發表一個機率型的配對模式，將消費者的意見產品屬性進行配對，以取得較佳的音樂產品推薦品質。由於混合式推薦模型，不僅保有內容式推薦與協同式推薦的優點，而且推薦品質亦較個別推薦模式來得較佳，因此本研究將採用混合式推薦模型，以推薦音樂 CD 給使用者。

### 3. 研究方法

#### 3.1 資訊推薦運算

由於音樂 CD 產品的種類眾多且複雜，使用者常無法快速找到想要的產品。目前的音樂推薦系統，僅記錄使用者所有交易資料以進行分析與推薦。因此，若能從交易資料中，找出購買相似產品的使用者群，並從中發現購買偏好，且針對購買偏好進行產品推薦，將能有效縮短使用者搜尋產品的時間。

本研究之資訊推薦的運算，是由關聯規則探勘 (association rules mining) (Agrawal and Srikant, 1994) 延伸而來。關聯規則探勘主要是找出產品之間的關係 (item association) 以及使用者之間的關係 (user associations)，並利用這兩種概念的互相搭配以做為推薦基礎。第一種為產品之間的關係，可以表現出哪些產品會時常一起購買，這些產品彼此之間可能隱含著某種關係而造成常常被一起購買的現象，因此當使用者購買其中某樣產品時，則可以推薦其它相關的產品給他。第二種為使用者之間的關係，找出購買習慣相似的使用者群組，因此要對群組中某個使用者進行推薦時，系統可以推薦相同群組中其他人時常購買的產品。例如，一群時常購買古典音樂 CD 的使用者，若相同群組的使用者購買音樂 CD，則可以推薦相同群組中其他人所購買過的古典音樂 CD。

在關聯規則探勘的運算上，又以 Apriori 運算式為代表，其運算首先要找出頻繁產品集合 (frequent item sets)，此集合的支持度必須大於使用者所訂定的最小支持度 (minimum support)；其次，是利用前一個步驟所產生的頻繁產品集合以歸納出關聯規則。由於 Apriori 運算式必須不斷的與資料庫資料進行運算，因此運算的效率相對較低。本研究則以此問題為出發點進行改善，作法主要是先將可能的頻繁產品集合先行聚焦，即先根據過去的銷售記錄，挑選出相關產品，進而推薦出 TOP-N 的清單(協同式推薦)，再依使用者偏好屬性進行篩選(內容式推薦)，並應用 Apriori 運算式對於資料裁剪的作法 (prune step)，在此則以配對係數做為資料裁減的門檻值，其主要的演算列如圖 1。

```

Algorithm SelectMusicCD( $L_k$ , userreq $_m$ , n)
input:  $C_k$ , userreq $_m$ , n
output:  $L_k$ 
global var: c.count
begin
     $C_n$  =transsort( $C_k$ , n);
    for all cadidates  $c \in C_n$  do begin
        map_coefficient( $c_j$ )= CalMapCoefficient( $c_j$ , userreq $_m$ );
        If (map_coefficient( $c_j$ ))>=map_lowbound
            insert  $c_j$  into  $L_k$ ;
        endif
        c.count++;
    end
end
end
    
```

圖 1 音樂 CD 推薦演算法

圖 1 的演算式表示：

1. 首先根據所有的交易紀錄( $C_k$ )，找出與目標產品( $C$  的前銷售量前  $n$  個頻繁產品集合( $C_n$ )。
2. 逐一比對與計算頻繁產品集合( $C_n$ )與使用者偏好(user request)的配對係數 (map\_coefficient)。配對係數的分析是透過使用者輸入偏好與協同式推薦結果做配對運算，計算方式首先是比對各產品之間的屬性值是否有相同屬性，比對屬性值的定義為0或1，若兩者都有的屬性則為1（兩者皆為1），兩者都沒有的屬性則為0（兩者皆為0），兩者的屬性不同則不列入計算（一者為1與一者為0），配對係數計算如下所示(周文賢, 2004)。

$$\text{Map\_coefficient} = \frac{a + b}{I} \quad , \quad (1)$$

$a$ ：使用者輸入產品 $n$ 與比對產品 $m$ 都有的屬性（兩者皆為1）總數，  
 $b$ ：使用者輸入產品 $n$ 與比對產品 $m$ 都沒有的屬性（兩者皆為0）總數，  
 $I$ ：所有屬性之總數。

3. 設定一配對係數的門檻值(map\_lowbound)，一旦前述的配對係數大於或等於門檻值，即列入推薦的內容( $L_k$ )。

此運算式因先前將頻繁產品集合先一次且大幅度的縮小，可以減少電腦運算時間。另外，由於先前的文獻探討中皆顯示，許多的資訊推薦系統皆必須先獲的使用者偏好資料，才能進行後續的運算。本文則讓使用者自主性的進行 CD 偏好的屬性挑選，再進行與原來建立的頻繁產品集合進行配對，在做法上可以避免前備資料的取得困難問題，因

而此推薦法在實務上應為可行。

### 3.2 CD 產品資料收集

為標示所有在音樂 CD 中之關鍵詞，本研究擷取「Joyaudio」網站中的音樂 CD 資訊做為擷取關鍵詞之資料，而關鍵詞的選取是依據音樂 CD 資訊中的字詞作分類的線索，選取標準為此關鍵字在該音樂 CD 資訊中出現，利用人工瀏覽的方式，將屬於各類別的關鍵詞逐一蒐集。

CD 產品資料分類描述如圖 2 所示，每一種音樂 CD 都有許多不同的因素，而 CD 產品是由許多產品屬性所組成。吳佳幸（2003）提出構成音樂的六大要素為節奏（rhythm）、結構（texture）、旋律（melody）曲式（form）、音色（tone color）與調性（tonality）。在本個案中產品因素採用音樂 CD 曲式、音色與特色；音樂 CD 的曲式屬性是依據「Joyaudio」網站所有音樂 CD 的分類，包含交響樂、電子音樂、流行音樂、爵士音樂、鄉村音樂、搖滾音樂、世界音樂、古典音樂、新世紀音樂與其他音樂十種屬性。音樂 CD 音色屬性是由本研究分析萃取出音樂 CD 資訊中西洋樂器之關鍵詞，包括鍵盤類、撥弦類、木管類、銅管類、弓弦類、打擊樂器類、人聲類與其他類八種屬性。音樂 CD 特色屬性也是由本研究分析擷取出音樂 CD 資訊中音樂旋律之關鍵字，包括輕快、抒情、浪漫、溫柔、悲傷、清晰、即興、現場、熱情與其他十種，而此十種關鍵詞屬性是從同義關鍵詞中挑選一個代表關鍵詞，例如，關鍵詞輕快，其同義詞共有十一個，分別為輕鬆、舒暢、快樂、欣喜、喜悅、愉悅、激揚、輕巧、歡悅、飽滿與舒服，因此取輕快為代表關鍵詞，其餘依此類推，如表 1 所示。

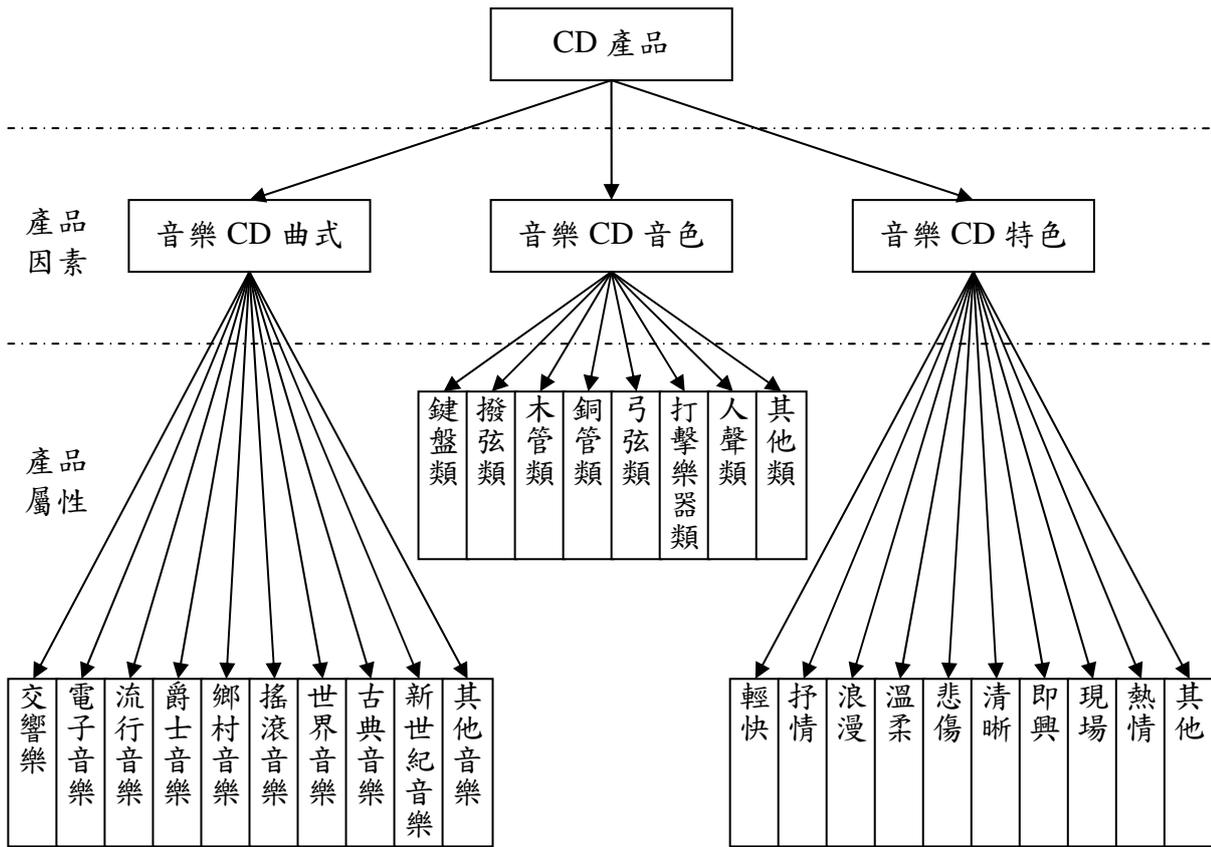


圖 2 音樂 CD 資料分類

表 1 音樂 CD 特色屬性表

關鍵詞	同義詞
輕快	輕鬆、舒暢、快樂、欣喜、喜悅、愉悅、激揚、輕巧、歡悅、飽滿、舒服
抒情	忘情
浪漫	舒適、情意綿綿
溫柔	柔和、柔順、柔美、低聲、輕聲、低沉、寧靜、靜謐、圓潤、溫暖、溫潤、悠揚、甜美、溫馨、柔情、動聽、豐潤、甜蜜、沉吟、溫潤、溫和
悲傷	淒涼、悲苦、悲戚、鬱悶、潺潺、銷魂、憂鬱、陰沉、淒美
清晰	明亮、清爽、清楚、清澈、乾淨、清新、純淨
即興	自由發揮、隨興
現場	
熱情	狂熱、狂野

藉由上述 CD 產品屬性定義，本研究分析產品資訊中相關同義的關鍵字來做為產品屬性，對產品屬性值的定義為 0 或 1，若產品有此屬性則為 1，無此屬性則為 0。我們以下面的例子說明，分析產品內容屬性之過程，圖 3 為本個案中的音樂 CD 產品，本研究分析重要關鍵字，從其中挑出爵士、鋼琴、貝斯、鼓、輕快、清晰與即興等關鍵詞如圖

3 粗體所標示；記錄方式如表 2，由表 2 得知產品屬性在音樂 CD 曲式是屬於爵士音樂，音色為鍵盤、撥弦和打擊樂器，特色是輕快、清晰和即興，將擷取出來的產品屬性記錄下來儲存至系統資料庫。

簡介		<p style="text-align: center;">爵士巴哈 Jacques Loussier Plays Bach</p> <p style="text-align: center;">爵士三重奏</p> <p>賈克·路西耶爵士樂三重奏 (賈克·路西耶 鋼琴/ 賽戈札克 貝斯/ 阿爾皮諾 鼓)</p> <p>爵士樂 20 世紀中期以獨特的魅力風行全球，其原因在於自由輕鬆的氣氛和無拘無束的即興演奏風格，而巴哈本身是個偉大的即興作曲家，他綿長的旋律線和完美的對位手法，使得不少喜歡「搖擺古典」的音樂家鍾情不已，紛紛拿巴哈的作品來改編。</p> <p>這張專輯中的 11 首名曲全部是他親自改編以三重奏型式演出。三個人在完美的默契下，展現各自的功力與獨奏技巧聽起來十分自然動聽。錄音方面清晰而平衡，仔細聆聽 Bass 的話，會有很不同的聽覺享受哦！</p> <p style="text-align: center;">(資料來源：<a href="http://www.joyaudio.com.tw/">http://www.joyaudio.com.tw/</a>)</p>
----	---	---

圖 3 音樂 CD 產品內容範例

表 2 音樂 CD 屬性表

屬性 CD	音樂 CD 曲式										音樂 CD 音色						音樂 CD 特色											
	交響樂	電子音樂	流行音樂	爵士音樂	鄉村音樂	搖滾音樂	世界音樂	古典音樂	新世紀音樂	其他音樂	鍵盤類	撥弦類	木管類	銅管類	弓弦類	打擊樂器類	人聲類	其他類	輕快	抒情	浪漫	溫柔	悲傷	清晰	即興	現場	熱情	其他
音樂 CD 屬性	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	1	1	0	0	0

藉由 CD 產品屬性定義，由本研究分析擷取出音樂 CD 資訊中同義關鍵字做為產品屬性，而且將產品屬性量化並記錄下來。

#### 4. 實證分析

本文目的是以混合式資訊推薦機制來協助使用者快速找到需要音樂 CD，本研究以「Joyaudio」發燒片之音樂 CD 網站為例，此網站的實體店面成立於 1999 年，主要是進口一些優質音樂 CD 產品，如 CD 與黑膠唱片等，客源以追求高品質音樂的客戶居多，

並且十分注重與顧客的溝通以及音樂經驗分享。

#### 4.1 協同式推薦步驟

##### 1. 從交易記錄中挑選 CD 產品類型

系統依過去交易記錄中，篩選出與 CD 產品相關記錄來進行分析。本文以產品 CD6 為使用者選購的 CD 產品為例，列如表 3，可得知系統會挑選出與產品 CD6 相關推薦記錄，且經由分析得到選購產品 CD6 時，也有可能選購產品 CD2、CD51、CD75、CD81、CD98 與 CD102。

表 3 CD 產品篩選記錄表

使用者	交易項目
U <sub>9</sub>	CD2、CD6、CD75、CD81
U <sub>14</sub>	CD2、CD6、CD51、CD81
U <sub>38</sub>	CD6、CD51、CD75、CD98
U <sub>55</sub>	CD2、CD6、CD51、CD75
U <sub>68</sub>	CD6、CD51、CD75、CD81
U <sub>78</sub>	CD6、CD51、CD98、CD102
U <sub>82</sub>	CD2、CD6、CD51、CD75
U <sub>90</sub>	CD6、CD75、CD81、CD102
U <sub>99</sub>	CD2、CD6、CD51、CD98

##### 2. 計算 CD 產品推薦次數與排序

首先必須計算 CD 產品推薦次數與排序，再進行 TOP-N 的產品之推薦。從上述表 8 中，分析得到每項音樂 CD 被推薦次數如表 4，可得知產品 CD2、CD51、CD75、CD81、CD98、CD102 被推薦次數為 5、7、6、4、3、2；經由排序可得到產品 CD2、CD51、CD75、CD81、CD98、CD102 的推薦排名。

表 4 產品推薦次數與排序表

產品名稱	推薦次數	排序
CD2	5	3
CD51	7	1
CD75	6	2
CD81	4	4
CD98	3	5
CD102	2	6

### 3. 推薦 TOP-N 項音樂 CD

使用者選購產品時，系統會依照協同式法則推薦 TOP-N 的產品清單給使用者參考如表 4，可得知選購產品 CD6 時，系統會推薦產品 CD51、CD75、CD2、CD81、CD98 與 CD102。

### 4. 採用推薦音樂 CD

若使用者接受上述協同式推薦之產品清單，則直接採用，若不採用，可以再利用內容式推薦，以挑選出更符合使用者需要的音樂 CD 產品。

## 4.2 內容式推薦步驟

### 1. 依使用者偏好屬性篩選協同式推薦音樂 CD

使用者也可以輸入偏好屬性如表 5 所示，使用者想要進一步找尋有關爵士音樂，在音色上可以是鍵盤，也可以是撥弦或打擊樂器，而音色是輕快、清晰、即興、熱情皆可，再與協同式推薦音樂 CD 產品屬性進行篩選，以挑選出想要的產品如表 5，可得知共篩選出 4 項音樂 CD 分別為 CD51、CD75、CD98 與 CD102，並對挑選出來的產品進行配對係數之計算。

表 5 使用者輸入與協同式推薦產品屬性表

屬 性 CD	音樂 CD 曲式										音樂 CD 音色						音樂 CD 特色											
	交響音樂	電子音樂	流行音樂	爵士音樂	鄉村音樂	搖滾音樂	世界音樂	古典音樂	新世紀音樂	其他音樂	鍵盤類	撥弦類	木管類	銅管類	弓弦類	打擊樂器類	人聲類	其他類	輕快	抒情	浪漫	溫柔	悲傷	清晰	即興	現場	熱情	其他
使用者輸入的屬性	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	1	1	0	1	0
CD51	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	0	0
CD75	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	1	0	1	0	0	0	0
CD98	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0
CD102	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	0	1	0	0	1	0	0	0

### 2. 配對係數計算

配對係數的分析是透過使用者輸入其偏好再將協同式推薦之產品做配對係數的計算，計算方式是將兩個產品之間的屬性進行比對如表 6 所示，共篩選出 4 項音樂 CD 產品分別為 CD51、CD75、CD98 與 CD102，再與使用者輸入的屬性進行比對，其配對係數值分別為 86%、68%、79% 與 75%，最後推薦係數值 > 75% (預設值) 的音樂 CD 產品。

表 6 使用者輸入與內容式推薦產品屬性之配對係數計算表

屬性 CD	音樂 CD 曲式									音樂 CD 音色							音樂 CD 特色							配對係數					
	交響樂	電子音樂	流行音樂	爵士音樂	鄉村音樂	搖滾音樂	世界音樂	古典音樂	新世紀音樂	其他音樂	鍵盤類	撥弦類	木管類	銅管類	弓弦類	打擊樂器類	人聲類	其他類	輕快	抒情	浪漫	溫柔	悲傷		清晰	即興	現場	熱情	其他
使用者輸入的屬性	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	1	1	0	1	0	
CD51	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	0	0	86%
CD75	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	1	0	1	0	0	0	0	68%
CD98	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	79%
CD102	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	75%

3. 推薦係數 ≥ 門檻預設值之音樂 CD

如上述所述，我們選擇配對係數值 ≥ 75% (預設值) 產品來進行推薦，如表 7 所示，可知推薦產品係數值 ≥ 75% 共有 3 項音樂 CD 產品分別為 CD51、CD98 與 CD102，係數值分別為 86%、79% 與 75%，由於此三項音樂 CD 產品與使用者偏好看比較相近，因此推薦給使用者參考。

表 7 推薦音樂 CD 產品

屬性 CD	音樂 CD 曲式									音樂 CD 音色							音樂 CD 特色							配對係數					
	交響樂	電子音樂	流行音樂	爵士音樂	鄉村音樂	搖滾音樂	世界音樂	古典音樂	新世紀音樂	其他音樂	鍵盤類	撥弦類	木管類	銅管類	弓弦類	打擊樂器類	人聲類	其他類	輕快	抒情	浪漫	溫柔	悲傷		清晰	即興	現場	熱情	其他
使用者輸入的偏好	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	1	1	0	1	0	
CD51	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	0	0	86%
CD98	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	79%
CD102	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	75%

## 5. 雛型系統說明

為了評估研究所提出之推薦方法的可行性，本研究開發一個雛型系統，首先介紹系統架構，接下來說明資料庫之建置，最後介紹系統的使用流程，並且展示各個操作畫面。

### 5.1 系統設計架構

本系統採用 three-tier 架構進行開發，在資料庫方面採用 Microsoft Access 2003 資料庫系統，應用程式方面則以 ASP 撰寫，撰寫內容除了系統界面，最後再利用 Internet Information Services 伺服器與使用者進行互動溝通，系統架構圖列如圖 4。

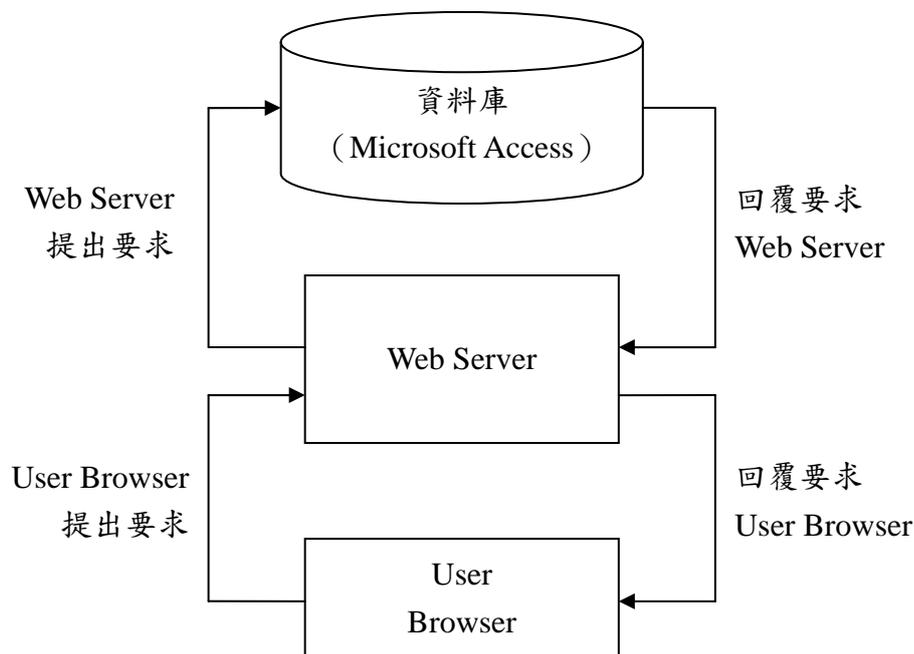


圖 4 系統架構圖

CD 資料庫的內容包含名稱、編號、圖片、簡介及其相關屬性；而屬性是透過人工方式來擷取分析音樂 CD 資訊中相關同義的關鍵字來做為產品屬性，對產品屬性值的定義為 0 或 1，若產品有此屬性則為 1，無此屬性則為 0，儲存於 Access 資料庫中。

### 5.2 系統操作流程

系統操作流程如圖 5 所示，接下來分別說明每一個流程的內容。

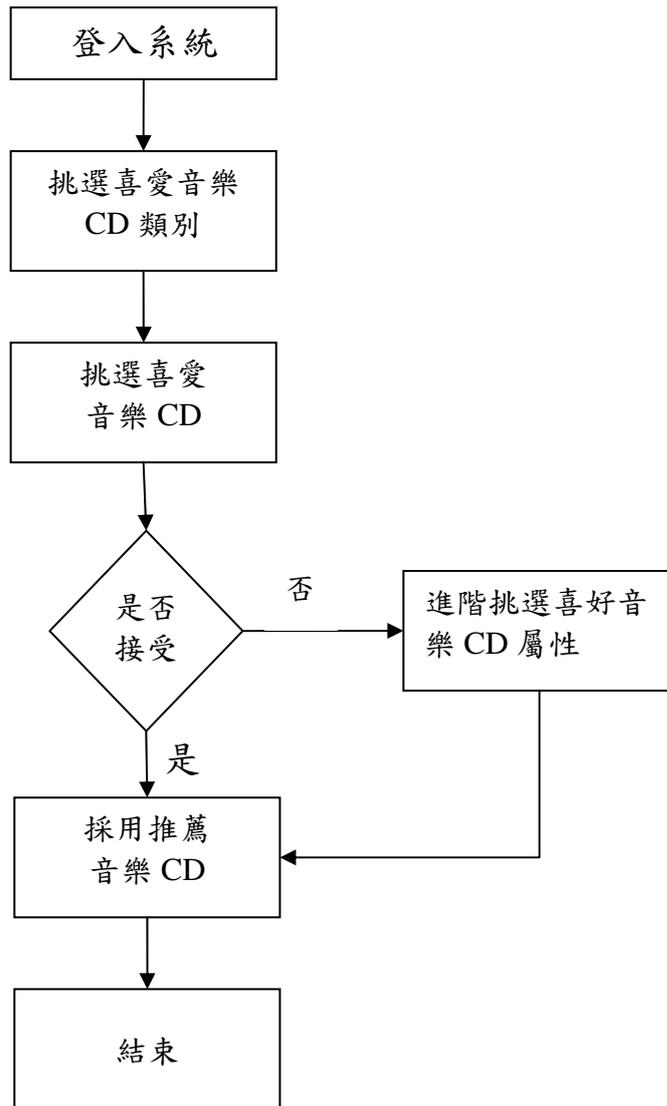


圖 5 系統使用流程圖

首先，使用者必須先行登入系統，系統展示的 CD 將分門別類展示，使用者須先挑選 CD 的類別後再挑選 CD。例如，當使用者點選爵士音樂類之音樂 CD 類別，則會進一步的顯示出該音樂 CD 類別中所有音樂 CD，供使用者瀏覽，其離型系統畫面如圖 6 所示。



圖 6 音樂 CD 類別瀏覽-爵士音樂類  
(CD 資料來源：<http://www.joyaudio.com.tw/>)

當使用者點選爵士女伶 5 時，協同式推薦系統會主動依過去的銷售記錄，進行分析以找出其可能有興趣的相關音樂 CD 供使用者做為瀏覽之參考，因此系統推薦 6 張音樂 CD，其系統畫面如圖 7 所示。



圖 7 協同式音樂 CD 推薦-爵士音樂類  
(CD 資料來源：<http://www.joyaudio.com.tw/>)

若使用者不接受或覺得上述系統所推薦的音樂 CD 數量太多，則可藉由內容式推薦的進階篩選使用者喜好音樂 CD 屬性，以過濾協同式推薦之 CD 產品，並計算兩者屬性之配對係數，如使用者喜愛音樂曲式為爵士音樂，音樂音色為撥弦類與人聲類，音樂特色為溫柔與清晰，最後系統推薦預設篩選門檻值為 75% 的音樂 CD，因此系統共推薦三張音樂 CD，其相似度分別為 93%、93% 與 86%，其雛型系統畫面如圖 8 所示。



圖 8 推薦篩選門檻值為 75%(預設值)的音樂 CD

(CD 資料來源：<http://www.joyaudio.com.tw/>)

## 6. 結論

近年來由於電子商務的蓬勃發展，造成產品資訊的負載，由於音樂 CD 產品的種類繁多且複雜的屬性，想要從眾多複雜產品中，找到符合使用者的產品是相當不容易的，因此各種產品推薦系統相繼被提出。本研究以混合式資訊推薦方法進行音樂 CD 資訊的推薦工作，嘗試找出符合使用者偏好的音樂 CD 產品。

本研究在做法上先利用協同式推薦機制，使用者進入系統選購 CD 產品時，系統會根據過去交易記錄中，挑選出與 CD 產品相關的產品記錄，並分析音樂 CD 屬性和計算 CD 產品與排序推薦次數，最後推薦 TOP-N 的產品清單提供使用者參考。若使用者不接受協同式推薦產品，再利用內容式推薦機制，由使用者輸入其偏好的產品屬性來篩選協同式推薦產品與配對係數之計算，最後系統推薦篩選門檻值以上的音樂 CD 產品給使用者參考。

本研究以「Joyaudio」網站為例，資料分析結果顯示，混合式資訊推薦方法在找尋音樂 CD 資訊上獲得良好的成效。本研究後續的發展可嘗試增加影響協同式與內容式推薦的相關屬性，並收集更多的使用者與產品資料，以更精確推薦出使用者真正想要的音樂 CD 產品。

## 參考文獻

1. 卜小蝶(1998), 「淺析個人化服務技術的發展趨勢對圖書館的影響」, 國立成功大學圖書館館刊, 第二期, 63-73 頁。
2. 吳佳幸(2003), 「以使用者為中心的音樂推薦系統」, 國立屏東科技大學資管所碩士論文。
3. 周文賢(2004), 「多變量統計分析 SAS / STAT 使用方法」, 第一版, 台北: 智勝文化事業有限公司。
4. Agrawal, R. and Srikant, R. (1994), "Fast Algorithm for Mining Association Rules," *In Proceedings of the VLDB Conference*, pp. 487-499.
5. Armstrong, R., Freitag, D., Joachims, T., and Mitchell, T. (1995), "WebWatcher: A Learning Apprentice for the World Wide Web," *In Proceedings of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments, Stanford*, pp. 6-12, CA.
6. Balabanovic, M. and Shoham, Y. (1997), "Fab: Content-Based, Collaborative Recommendation," *Communications of the ACM*, 40(3), pp. 66-72.
7. Cao, Y. and Li, Y. (2007), "An Intelligent Fuzzy-based Recommendation System for Consumer Electronic Products," *Expert Systems with Applications*, 33(1), pp. 230-240.
8. Chai, W. and Vercoe, B. (2000), "Using User Models in Music Information Retrieval Systems," *In Proceedings of the International Symposium on Music Information Retrieval*, pp. 201-206.
9. Chao, Y. H. and Kim, J. K. (2004), "Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendations in E-Commerce," *Expert Systems with Applications*, 26(2), pp. 233-246.
10. Chen, H. C. and Chen, A. L. P. (2001), "A Music Recommendation System Based on Music Data Grouping and User Interests," *In Proceedings of the ACM International Conference on Information and Knowledge Management CIMK'01*, (5-10), pp. 231-238.
11. Choi, S. H., Kang, S., and Jeon, Y. J. (2006), "Personalized Recommendation System Based on Product Specification values," *Expert Systems with Applications*, 31(3), pp. 607-616.
12. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., and Sartin, M. (1999), "Combining Content-Based and Collaborative Filters in an Online Newspaper," *In Proceedings of the ACM SIGIR Workshop on Recommender Systems*.
13. Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992), "Using Collaborative Filtering to Weave an Information Tapestry," *Communications of the ACM*, 35(12), pp. 61-70.
14. Hayes, C. and Cunningham, P. (2004), "Context Boosting Collaborative Recommendations," *Knowledge-Based Systems*, 17(2-4), pp. 131-138.

15. <http://www.joyaudio.com.tw> °
16. Kim, B. D. and Kim, S. O. (2001), "A New Recommender System to Combine Content-Based and Collaborative Filtering Systems," *Journal of Database Marketing*, 8(3), pp. 244-252.
17. Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. (1997), "GroupLens: Applying Collaborative Filtering to Usenet News," *Communications of the ACM*, 40(3), pp. 77-87.
18. Krulwich, B. and Burkey, C. (1997), "The InfoFinder Agent: Learning User Interests through Heuristic Phrase Extradion," *IEEE Intelligent Systems Journal (Expert)*, 12(5), pp. 22-27.
19. Lang, K. (1995), "NewsWeeder: Learning to Filter Netnews," *In Proceedings of the 12th International Conference on Machine Learning*, pp. 331-339.
20. Lawrence, R. D., Almasi, G. S., Kotlyar, V., Viveros, M. S., and Duri, S. S. (2001), "Personalization of Supermarket Product Recommendations," *Data Mining and Knowledge Discovery*, 5(1-2), pp. 11-32.
21. Li, Q., Myaeng, S. H., and Kim, B. M. (2007), "A Probability Music Recommender Considering User Opinions and Audio Features," *Information Processing and Management*, 43(2), pp. 473-487.
22. Liu, D. R. and Shih, Y. Y. (2005), "Hybrid Approaches to Product Recommendation Based on Customer Lifetime value and Purchase Preferences," *The Journal of Systems and Software*, 77(2), pp. 181-191.
23. Roh, T. H., Oh, K. J., and Han, I. (2003), "The Collaborative Filtering Recommendation Based on SOM Cluster-Indexing CBR," *Expert Systems with Application*, 25(3), pp. 413-423.
24. Shardanand, U. and Maes, P. (1995), "Social Information Filtering: Algorithms for Automating "Word of Mouth" ," *In Proceedings of the CHI'95 Conference on Human Factors in Computing System*, pp. 210-217.
25. Terveen, L., Hill, W., Amento, B., McDonald, D., and Creter, J. (1997), "PHOAKS: a System for Sharing Recommendations," *Communications of the ACM*, 40(3), pp. 59-62.
26. Wang, F. H. and Shao, H. M. (2004), "Effective Personalized Recommendation Based on Time-Framed Navigation Clustering and Association Mining," *Expert Systems with Applications*, 27, pp. 365-377.
27. Wasfi, A. M. A. (1999), "Collecting User Access Patterns for Building User Profiles and Collaborative Filtering," *In Proceedings of the 1999 International Conference on Intelligent User Interfaces*, pp. 57-64.
28. Zhang, Y. and Jiao, J. (2007), "An Associative Classification-based Recommendation System for Personalization in B2C E-commerce Applications," *Expert Systems with Applications*, 33(2), pp. 357-367.