

2010 TOPCO 崇越論文大賞

論文題目

基於 RFM 分析法之顧客適性化

產品推薦機制

報名編號： C0026

摘要

為了解顧客喜好並滿足其需求，企業經常仰賴推薦系統來提供符合顧客個人化之產品或服務。過去曾提出產品適時性(Timely)推薦方法，但是當產品遇到重覆週期的天數時，系統仍會予以推薦，因此容易造成重覆推薦因而降低推薦效率。爾後亦有產品週期性推薦，主要以產品被購買的最小與最大天數區間作為推薦顧客產品之依據，不過這些研究忽略了低忠誠度顧客對產品可能有興趣但卻未購買的潛在偏好之重要性。有鑑於此，本研究考慮產品購買週期與顧客消費特性，提出一套基於 RFM 分析法之顧客適性化產品推薦機制。由實驗結果顯示本研究之推薦機制能提供更有效益的產品推薦。

關鍵字：RFM 分析法、適性化、自組織映射圖網路、推薦系統

壹、緒論

一、背景

隨著網際網路技術與電子商務的蓬勃發展，企業為了增加利潤並提高產品銷售機會，必須改變傳統型態的行銷策略與手法。依據 80/20 法則，企業中 80% 的利潤是來自 20% 對企業具有忠誠度之顧客，由此可知具有忠誠度的顧客在企業經營裡的角色。企業如何得知顧客需求，以及顧客如何能取得企業的幫助，都必須依賴有效的資訊，為了能幫助顧客以快速又有效的方式搜尋出符合需求的資訊或服務，並輔助企業了解顧客喜好以滿足需求，推薦系統(Recommendation System, RS)便因應而生。

推薦系統是以資訊過濾(Information Filtering, IF)為基礎，且能利用個人化服務技術(Goldberg et al., 1992)來降低資料過載的問題，並篩選出顧客所需的資訊。其方式是藉由顧客的個人歷史紀錄來分析顧客的消費特性與喜好，也可透過具有相似偏好的其它顧客之歷史交易紀錄，預測並推論出符合顧客需求的相關產品，藉此主動並適時地提供顧客潛在的資訊或產品。

為了輔助推薦系統的開發，過去有許多學者提出不同的推薦方法，即以顧客面或資料面之分析找出某些特性或關聯，藉以作為推薦之依據，使推薦系統可廣泛地應用在不同領域並提升其推薦品質。顧客面之分析例如：顧客屬性、對產品或服務的評價、行動所在地等；資料面之分析包含：顧客歷史交易紀錄、網頁瀏覽路徑、產品類別或時間屬性，透過不同面向進一步分析與觀察其具有何種特性，並針對所分析出的特性推薦適合顧客的產品或服務，這種能針對使用者特性來做推薦的方式，本研究稱之為「適性化推薦」。

在過去相關的適性化推薦研究中，有學者(Yuan and Tsao, 2003)分析顧客之個人輪廓檔(User Profile)的屬性資料，再依據顧客目前所在的位置，推薦適地性且符合顧客喜好的廣告資訊；亦有學者(Kim et al., 2002)提出藉由顧客在網路上的瀏覽路徑與點選行為，推論相同類別下具有潛在偏好的产品並予以推薦；Min(2005)等學者則依據時間屬性觀察隨著不同的時間所變化的顧客行為，並分析群體偏好予以推薦；近年更有研究(Ha, 2007)觀察每一段時間區間內群體顧客之 RFM 值的變化，推論下一區間該群顧客的消費特性。

二、動機

由上述可知，過去有關推薦之研究多以顧客特性或對產品的評價分析做為推薦基礎，大部分的研究均未考慮消費性產品被購買的週期，故無法在適當的時間提出推薦，因此，就算推薦了對的產品給顧客，但顧客對該產品仍未達需求的時間，顧客未必會購買所推薦之產品。因此若能同時考慮顧客的產品偏好、個人化的消費習性與消費性產品被購買的週期，如此將可以更準確地推薦已達購買週期的產品給顧客，進而大幅提升推薦效能。

為近年來有學者(Li et al., 2006)提出適時性(Timely)的推薦方法，該研究針對高忠誠度顧客對某項產品的需求期間到達適時性的推薦，即分析顧客進入賣場或網站時對產品的需求狀況，但這種推薦是以天數而非產品被需求的區間做考量，因此當產品遇到重覆週期的天數時，可能會重覆推薦因而降低推薦效率。

為改善上述的問題，學者(Li et al., 2009)再提出一套具有消費性產品被購買週期的推薦系統，針對顧客之個人化消費習性與其對產品週期性的需求方式，以了解顧客對產品所具有之週期性購買關係並分析產品被購買區間，期望在有效的區間內主動推薦顧客產品。此方法主要是針對經常性消費的高忠誠度顧客予以分析並推薦，不過這個方法忽略了低忠誠度顧客對產品有潛在興趣偏好之重要性。因此若能準確掌握高忠誠度顧客之產品需求，同時發掘低忠誠度顧客之潛在產品偏好，如此便能有效提升推薦效能。

三、目的

有鑑於此，本研究欲提出一套基於 RFM 分析法之顧客適性化產品推薦機制，期望達到以下目的：(1)整合並改善過去學者未考慮產品重覆購買週期之缺失。(2)依據顧客之消費特性，分析顧客喜好與所需之產品並給予推薦。(3)分析高忠誠度顧客之產品購買週期，及早主動提供所需的產品，掌握顧客消費動向。(4)發掘低忠誠度顧客潛在偏好之產品，期望提升低忠誠度顧客之推薦效能，進而刺激低忠誠度顧客轉為高忠誠度顧客。

貳、文獻探討

一、RFM 分析法

隨著資訊科技的蓬勃發展，在商務盛行的社會中，提升推薦品質、滿足顧客需求並保有長期的競爭優勢才能創造最大的利潤，因此顧客終身價值(Customers Lifetime Value, CLV)便成了重要的一個因素(Liu and Shih, 2005)。顧客終身價值是一種代表顧客忠誠度之生命週期意義，為了分析顧客忠誠度，過去有學者(Liu and Shih, 2005)提出運用 RFM 分析法來衡量顧客對企業的忠誠度與貢獻度，RFM 是經常被用來作為市場區隔的方法之一(Schafer et al., 1999)。針對 R、F、M 這三個指標其詳細意義說明如下(Suh et al., 1999)：

- (一) 最近購買時間(Recency, R)：其定義為顧客最近一次消費與前次消費的時間距離，值越小即越接近現在時間，這是表示顧客越有可能重覆購買；而 R 指標與再次購買機率之間是為負向的關係。
- (二) 購買頻率(Frequency, F)：其定義為顧客於某段時間區間內購買產品的總次數，值越大即購買機率越高，表示顧客有越高的忠誠度；而 F 指標可用來代表顧客與企業間的互動關係，互動越高代表購買的可能性也越高，所以與再次購買的可能性之間是為正向的關係。
- (三) 購買金額(Monetary, M)：其定義為顧客於某段時間區間內購買產品的總金額，值越大即花費越大，表示顧客的需求越強烈；而 M 指標與再次購買機率之間是為正向的關係。

過去有許多 RFM 分析法相關的應用與研究，如 Hughes (1994)提出藉由計算 RFM 後所得之數值來區隔顧客，再針對不同群組中的顧客採用不同的行銷策略。更有學者(Liu and Shih, 2005)將 RFM 分析法結合於推薦系統，認為針對不同產品之特性，應給予 RFM 變數不同的權重值以求得更準確的顧客區隔結果，藉此瞭解顧客的消費習性與再次購買的機會，有效地針對目標顧客給予推薦。由此例可知，RFM 分析法能夠強化顧客個人的需求與喜好，更有助於個人化行銷策略的訂定，同時亦能增加顧客回購率。

二、週期性產品推薦方法

Li 等學者於 2009 年提出一套基於產品購買週期之個人化產品購買週期性推薦機制，強調顧客對產品應具有週期性的需求，並於某一時間點會產生週期性之消費行為。此方法第一步驟是計算顧客曾經購買過的產品類別 i 其所有交易時間差，第二步驟分析產品的單個消耗時間、最小購買週期與平均購買週期以瞭解顧客對產品的購買週期。最後第三步驟則是透過產品的購買週期誤差與週期標準差值之大小判斷，求得產品推薦週期區間 (Recommender Product Range,

PRP) $RPR_i[d_{\min}, d_{\max}]$ ，利用此推薦週期區間之最大值與最小值，即推薦的最大天數與最小天數，並根據顧客上一次購買時間作為推薦週期時間的基準點，由此預測出產品下一次被顧客購買的週期區間 $RPR_i[d_{\min} + T(i)_{\text{now}}, d_{\max} + T(i)_{\text{now}}]$ ，若顧客購買該類產品有落入此區間，表示預測該產品推薦週期成立，故推薦此產品給顧客。

過去有學者以 RFM 分析法、時間屬性或適時性推薦分析顧客的個人需求、喜好與瞭解其忠誠度，但這些方法對於產品被購買的週期與顧客重覆性需求並未加以重視，而產品購買週期之概念，主要是認為顧客對於不同產品應有其重覆性需求與其購買產品之週期性關係，期望瞭解顧客購買產品之週期性，於顧客對產品產生週期性需求前，適時、適切地提供顧客個人所需之產品，藉此掌握顧客的消費動向並提高個人化產品之推薦效能。有鑑於此，本研究將採用此方法來分析高忠誠度顧客過去曾經購買過的產品之週期區間，以作為產品推薦之依據。

三、推薦系統

推薦系統(Recommendation System, RS)係指一種具有過濾資訊機制的服務工具(Sarwar et al., 2000)，不僅可依據顧客的偏好或興趣，協助顧客解決資訊過載(Overload)的問題，亦可分析出顧客所需的資訊，並推薦符合顧客需求之產品或服務。目前推薦系統較常運用的機制為(Sarwar et al., 2000)內容導向式推薦(Content-Based Recommendation, CB)及協同過濾式推薦(Collaborative Filtering Recommendation, CF)，以下分別詳細說明：

- (a) 內容導向式：即依據顧客的個人輪廓檔(User Profile)，分析項目(Item)的屬性特徵並比對顧客興趣及瀏覽行為，以發掘顧客可能感興趣的項目並予以推薦。內容導向式推薦的優點是能快速且準確地推薦有關顧客個人偏好或興趣的產品。但此機制有以下三點限制：(1)對於媒體屬性項目如影像、圖片與聲音卻無法分析及判斷，因此降低了推薦項目的類型，(2)僅藉由個人輪廓中的資訊加以分析，無法發掘並推薦和顧客過去曾經購買過的其他相似產品，(3)無法辨識出相同性質但不同名稱的資訊，即容易有同義異名(Synonymy)的問題。
- (b) 協同過濾式：此方法是透過分析相似特性群組內成員的共同偏好或興趣，再計算顧客之間偏好行為的相似度，找出相似偏好的顧客並對其產生關聯性推薦。協同過濾式推薦是推薦其他顧客最可能喜好的前幾項產品(Top_N)給予目標顧客，藉由同群具有相似偏好的顧客來產生推薦項目，此方法能發掘目標顧客之潛在偏好，增加銷售機會。雖然協同過濾式推薦是目前在推薦系統中最常被使用的推薦機制，然而此方法也有其限制：(1)稀疏性(Sparisty)：意指顧客的交易資料過少導致推薦困難。(2)延展性(Scalability)：協同過濾式推薦是經由計算相似度，因此產品項目數量的大小或顧客人數的多寡，與計算複雜度呈正比關

係，可能降低推薦系統的整體效能。(3)冷啟始(Cold-start)：意指顧客或產品資訊不充足，導致系統難以計算顧客之間的相似程度，因而造成推薦或預測不足的問題。

由上述可知，內容導向推薦能快速準確地推薦符合顧客個人需求或偏好產品，因此較為適合做個人化的推薦。協同過濾式推薦則是利用群體之相似偏好推薦潛在產品給顧客，因此，本研究綜合上述二個方法，採用混合式推薦(Mixed Approach Recommendation)，擷取此兩種推薦機制的優點以期達到更好的推薦效能。

四、自組織映射圖網路(Self-Organizing Map, SOM)

自組織映射圖網路(SOM)屬於非監督式學習網路模式，是由 Tuevo Kohonen 在 1980 年提出，其基本原理可追溯自大腦中的腦神經細胞有「物以類聚」的特性。它可以將樣本空間中樣本點的聚類擷取出來，建構出二維平面的網路拓樸圖，能易於了解其聚類關係(葉怡成，2003)。自組織映射圖網路是一種競爭式網路並具有從輸入空間至群聚提供拓樸映射的特徵(Kohonen, 1995)。自組織映射圖網路採用鄰近區域(Neighborhood)的概念，鄰近區域內的輸出處理單元會相互影響，並透過競爭得勝之類神經元的鄰近區域內，來實現類神經元之間的交互作用。自組織映射網路架構如圖 1 所示。自組織映射圖網路包含以下特點：

- (一) 具空間視覺化：自組織映射圖網路的輸出單位，其相對位置具有意義，而一般類神經網路模式則無。因此，自組織映射網路是以二維空間的網路拓樸型式來呈現，可清楚、簡單觀察出資料聚類的分佈狀態。
- (二) 非監督式學習網路模型：非監督式學習網路具有從問題中取得訓練範例，並從中學習範例的聚類規則，以適用於新的案例，因此自組織映射圖網路於學習過程中不需訓練範例之相對目標值，即可有效的進行學習並呈現出聚類結果。

由於自組織映射圖網路的輸入資料型態為連續值，此方法既簡單又能快速地得出分群結果，並且能以拓樸圖像的方式清楚呈現資料之分佈情形，因此，在本研究中將以自組織映射圖網路作為低忠誠度顧客群之分群技術。

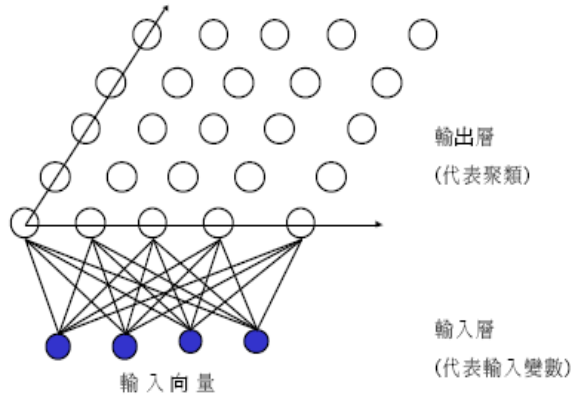


圖 1 自組織映射圖網路架構

五、評估指標

推薦系統的效能評估一般較常採用準確率(Precision)、回應率(Recall)以及 F1(Fallout)指標作為衡量推薦品質的指標(Oard and Marchionini, 1996)。所謂準確率(Precision)意指在系統推薦的產品項目中顧客實際購買之產品所佔的比率，而回應率(Recall)則代表顧客在實際購買之產品中系統所推薦之產品所佔的比率(參見公式(1)與公式(2))。然而，準確率與回應率之間存有不平衡的情形，即當系統推薦項目增加時，準確率因分母變大而會下降但回應率卻可能上升，因此有學者(Sarwar et al., 2000)將 Recall 與 Precision 結合提出 F1(Fallout)指標的衡量方法(參見公式(3))，藉由該指標協調兩者之間的矛盾，以便能更準確地衡量推薦之效能。本研究將以此三項評估指標來衡量推薦效能，以下分別說明各項評估指標之公式。

(一) Precision

$$Precision = \frac{Purchase_item \cap Recommend_item}{Recommend_item} \quad (1)$$

(二) Recall

$$Recall = \frac{Purchase_item \cap Recommend_item}{Purchase_item} \quad (2)$$

(三) F1

$$F1 = \frac{2 \times recall \times precision}{recall + precision} \quad (3)$$

參、研究方法

本研究基於產品購買之週期性與顧客之產品偏好，提出一套基於 RFM 分析法之顧客適性化產品推薦機制。即運用 RFM 分析法與自組織映射圖網路技術，考慮高忠誠度顧客於不同時間對不同產品具有週期性的需求，以及低忠誠度顧客對某

項產品之偏好作為本研究的推薦依據。研究方法包含五個步驟，以下將分別詳細說明各步驟之處理流程，流程架構如圖 2 所示。

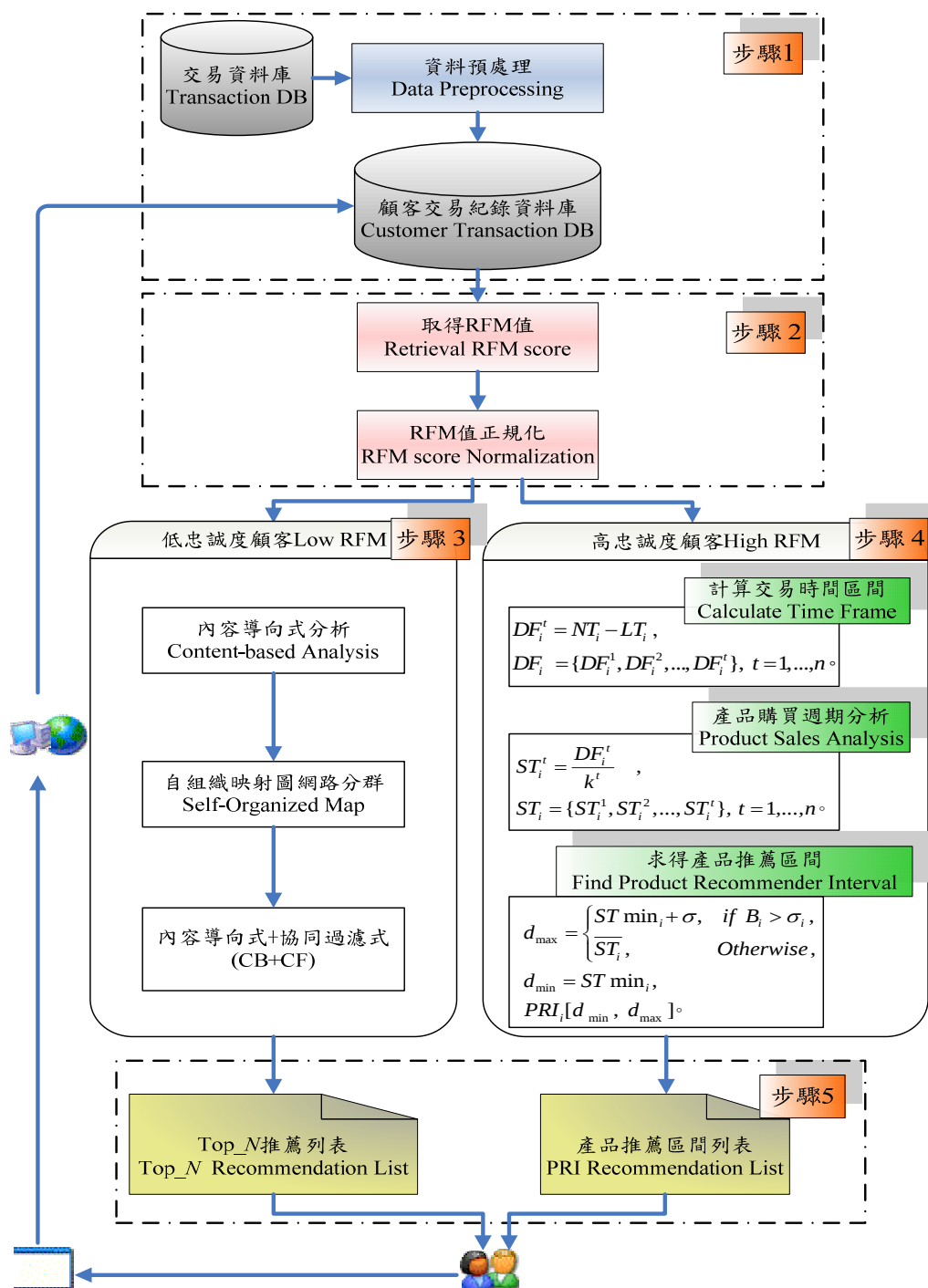


圖 2 研究架構

一、資料預處理

本研究是以消費性產品為研究內容，需透過大量的顧客交易紀錄及產品資料作為研究分析之實驗數據，所需資料是從交易資料庫(Transaction_Database, TDB)

中收集交易資料表(Transaction_Data_Table, TDT)、產品資料表(Product_Data_Table, PDT)、產品類別資料表(Product_Category_Data_Table, PCDT)、顧客資料表(Customer_Data_Table, CDT)與時間資料表(Date_Data_Table, DDT)，並利用關聯式查詢交叉彙整後取得每位顧客消費之交易紀錄，經篩選後剔除購買產品次數過於分散且無固定購買習性之顧客，以取得有用的分析資料，本研究為避免分析過程中發生稀疏性的問題，因此選擇產品類別(Product_category)作為產品分析項目。

二、顧客 RFM 分析

經過資料預處理步驟後，首先從顧客交易紀錄資料庫中取得每位顧客之 RFM 原始數值(OR_u 、 OF_u 、 OM_u)，此數值是針對每位顧客計算其最近一次消費時間與前一次消費時間差求得 R-value(OR_u)，再取得顧客於一年內的消費次數得到 F-value(OF_u)，最後將產品之銷售單價與顧客購買數量相乘求得 M-value(OM_u)。然而三個指標所代表的意義與數值範圍不盡相同，因此需先透過顧客五等分法(Customer Quintiles Method)將三個指標之數值做正規化的處理，亦即將 R-value、F-value 與 M-value 依據顧客忠誠度與貢獻度以 1 到 5 的分數(CR_u 、 CF_u 、 CM_u)表示之，並加總三個指標之數值藉此得到每位顧客的 RFM 總分($Score_u$)，此 RFM 總分即用以判斷該顧客忠誠度高低之依據。此五等分之計算步驟如公式(4)至公式(6)：

$$\bar{R} = \frac{R_h - R_l}{5}, R_l = \min\{R_1, R_2, \dots, R_n\}, R_h = \max\{R_1, R_2, \dots, R_n\} \quad (4)$$

$$\bar{F} = \frac{F_h - F_l}{5}, F_l = \min\{F_1, F_2, \dots, F_n\}, F_h = \max\{F_1, F_2, \dots, F_n\} \quad (5)$$

$$\bar{M} = \frac{M_h - M_l}{5}, M_l = \min\{M_1, M_2, \dots, M_n\}, M_h = \max\{M_1, M_2, \dots, M_n\} \quad (6)$$

$\bar{R}$: The average of all recency value, R

$\bar{F}$: The average of all frequency value, F

$\bar{M}$: The average of all monetary value, M

R_l : Minimum R-value

F_l : Minimum F-value

M_l : Minimum M-value

R_h : Maximum R-value

F_h : Maximum F-value

M_h : Maximum M-value

表 1 RFM 五等分數值對照表

value score	R-value(OR_u)	F-value(OF_u)	M-value(OM_u)
1	$R_h \sim \bar{R} \times 4$	$F_l \sim \bar{F} \times 1$	$M_l \sim \bar{M} \times 1$
2	$\bar{R} \times 4 \sim \bar{R} \times 3$	$\bar{F} \times 1 \sim \bar{F} \times 2$	$\bar{M} \times 1 \sim \bar{M} \times 2$
3	$\bar{R} \times 3 \sim \bar{R} \times 2$	$\bar{F} \times 2 \sim \bar{F} \times 3$	$\bar{M} \times 2 \sim \bar{M} \times 3$
4	$\bar{R} \times 2 \sim \bar{R} \times 1$	$\bar{F} \times 3 \sim \bar{F} \times 4$	$\bar{M} \times 3 \sim \bar{M} \times 4$
5	$\bar{R} \times 1 \sim R_l$	$\bar{F} \times 4 \sim F_h$	$\bar{M} \times 4 \sim M_h$

經由各指標之平均值，即 $\bar{R}$ 、 $\bar{F}$ 與 $\bar{M}$ ，可依據平均值之倍數將其範圍劃分為五等分，分別以 1 到 5 的數值(越大表忠誠度越高)表示，對照結果如表 1。由於 R 值的定義為顧客最近一次消費與前次消費的時間距離，R 值越小代表顧客對企業之忠誠度越高，因此將五等分轉換後 R 值越小分數越高，而 F 值與 M 值則相反。接著利用公式(7)將每位顧客之三項數值加總計算出新的 RFM 總分即($Score_u$)。例如顧客 u 之 $CR_u=4$ 、 $CF_u=3$ 、 $CM_u=5$ ，則 $Score_u=4+3+5=12$ ，並求出所有顧客之平均 RFM 值($\overline{Score}$)。透過公式(8)判斷若 $Score_u$ 高於 $\overline{Score}$ 表示為高忠誠度顧客群 RFM_{high} ，反之則為低忠誠度顧客群 RFM_{low} 。

$$\overline{Score} = \frac{\sum_{u=1}^n Score_u}{n} \quad (7)$$

$$\begin{aligned} &\text{if } Score_u < \overline{Score} \text{ then } u \in RFM_{low} \\ &\text{else } u \in RFM_{high} \end{aligned} \quad (8)$$

$Score_u$: The total RFM value of customer u

$\overline{Score}$: The average RFM value of all customers

n : Total number of customers

由於被分入到低忠誠度顧客群之顧客其交易頻率較低，故較難產生固定的產品購買頻率，因此本研究利用協同過濾式推薦，為低忠誠度顧客找出同類群之顧客喜好排行榜(Top_N)產品，並結合內容導向式推薦之顧客個人偏好產品，運用混合式推薦作為該群顧客之推薦依據。

三、自組織映射圖網路分群

上一階段已將所有顧客依據其消費特性分為高忠誠度顧客群與低忠誠度顧客群，由於低忠誠度顧客群較無固定購買習性，因此無法計算其產品購買需求，為了解低忠誠度顧客群對產品之潛在偏好，在此階段本研究運用自組織映射圖網路(Self-Organizing Map, SOM)進行分群，將具有相似消費行為之顧客形成同一群

組，藉以分析同群顧客之喜好排行榜(Top_N)，發掘顧客之潛在偏好產品並予以推薦，期望運用此機制來提升低忠誠度顧客群之推薦成效。

分群的方法是取得由上一階段分析出之顧客 RFM 五等分，以每位顧客交易紀錄之 CR_u 、 CF_u 、 CM_u 值作為自組織映射圖網路的輸入值，即輸入單元為 $I_u = \{X_{u1}, X_{u2}, X_{u3}\}$ ，其中 X_{u1} 為顧客 u 之 CR 值集合， $X_{u1} = \{CR_1, CR_2, \dots, CR_n\}$ ； X_{u2} 為顧客 u 之 CF 值集合， $X_{u2} = \{CF_1, CF_2, \dots, CF_n\}$ ； X_{u3} 則為顧客 u 之 CM 值集合， $X_{u3} = \{CM_1, CM_2, \dots, CM_n\}$ ，將上述變數分別送到 SOM 網路後即可得出網路之群聚結果，並產生輸出單元 $O = \{Y_1, Y_2, \dots, Y_m\}$ ，每一個 Y_m 代表一個群組且各群組內皆包含 1 至數位顧客，被分至同一群組內之顧客即代表這些顧客具有相似的消費特性。

接著，依據分群結果取得同群所有顧客之交易紀錄，採用協同過濾式(CF)推薦機制對每一群組予以分析該群顧客之喜好排行榜產品，經由計算該群顧客所購買的各類產品次數，次數越多代表該類產品對該群顧客越具有顯著偏好，因此會將該類產品列為推薦項目。

四、產品推薦區間分析

此階段係利用週期性方法(Li et al., 2009)，針對高忠誠度顧客群過去之歷史交易紀錄對其進行個人化產品購買需求之分析。由於高忠誠度的顧客具有購買頻率高且較固定之購買週期，因此本研究採用 Recency 與 Frequency 作為顧客個人產品推薦區間分析依據，預測顧客下一個週期可能發生的消費行為。其步驟如下：

步驟 1：計算產品之消費時間區間

此階段利用 R-value 之概念計算出每位顧客其 m 筆交易紀錄，顧客於第 t 次購買產品類別 i 之消費時間 NT_i (Now Time for Category i , NT_i) 距離前一次消費時間 LT_i (Last Time for Category i , LT_i) 的時間區間(天數) DF_i^t (Date Frame, DF_i^t)，並將求得之時間區間依序存入消費時間區間集合 DF_i 中，如公式(9)。

$$\begin{aligned} DF_i^t &= NT_i - LT_i, \\ DF_i &= \{DF_i^1, DF_i^2, \dots, DF_i^t\}, t=1, \dots, T. \end{aligned} \quad (9)$$

步驟 2：求得顧客購買產品之週期

首先需針對顧客過去曾購買過的產品計算各類產品的平均消費時間 ST_i^t 。即利用顧客於第 t 次交易購買產品類別 i 的 k^t 個數量，計算此次 i 類產品之單個消費時間 ST_i^t (Sales Time for Product Category i , ST_i^t)，所有計算出之 ST_i^t 依序放入產品消

費週期集合 $ST_i = \{ST_i^1, ST_i^2, \dots, ST_i^t\}$ 內，如 公式(10)，並可於消費週期集合 ST_i 中取得最小之購買週期 $ST \min_i$ (Minimal Sales Time, $ST \min_i$)，如 公式(11)所示，最後利用 公式(12) 計算出顧客對各產品類別 i 之平均購買週期 $\overline{ST_i}$ (Average Sales Time, $\overline{ST_i}$)。

$$ST_i^t = \frac{DF_i^t}{k^t}, ST_i = \{ST_i^1, ST_i^2, \dots, ST_i^t\}, t = 1, \dots, T. \quad (10)$$

$$ST \min_i = \min\{ST_i^t\}, t = 1, \dots, T \quad (11)$$

$$\overline{ST_i} = \frac{\sum_{t=1}^T ST_i^t}{T} \quad (12)$$

步驟 3：分析產品推薦區間

經過上述兩個步驟的計算後，此階段將分析出各產品類別 i 之產品推薦區間 (Product Recommender Interval for Product Category i , $PRI_i[d_{\min}, d_{\max}]$)。為能確實分析顧客對不同產品類別的推薦區間，必須先透過 公式(13) 計算平均購買週期與最小購買週期之間的週期誤差 B_i (Bias, B_i) 以及週期標準差 σ_i ，如 公式(14)。

$$B_i = \overline{ST_i} - ST \min_i \quad (13)$$

$$\sigma_i = \sqrt{\frac{1}{T} \sum_{t=1}^T (ST_i^t - \overline{ST_i})^2} \quad (14)$$

接著，依據週期誤差 B_i 與週期標準差 σ_i 大小之判斷，求得產品推薦區間 $PRI_i[d_{\min}, d_{\max}]$ 的最大推薦天數 $d_{\max}$ ，如 公式(15)，最小推薦天數 $d_{\min}$ 則為最小購買週期 $ST \min_i$ ，以產生此產品推薦區間 $PRI_i[d_{\min}, d_{\max}]$ 作為推薦顧客產品之週期依據。

$$d_{\max} = \begin{cases} ST \min_i + \sigma, & B_i > \sigma_i, \\ \overline{ST_i}, & \text{if} \\ \end{cases} \quad \text{Otherwise,} \quad (15)$$

$$d_{\min} = ST \min_i.$$

五、顧客適性化推薦列表

本小節在說明如何分別挑選出適合高忠誠度顧客與低忠誠度顧客之產品並適時地推薦給顧客。本研究之推薦列表是針對不同特性顧客以不同的方法所產生。

就低忠誠度顧客而言，其推薦列表是透過混合式推薦機制，將自組織映射圖網路分群後同群顧客具有偏好之排行榜(Top_N)產品，再結合顧客個人對產品之偏好，建立屬於該顧客之推薦列表給予顧客；對於高忠誠度顧客群，則依據顧客歷史交易紀錄分析顧客購買產品之週期，計算出顧客下一次對產品產生需求之產品推薦區間，若產品落入推薦區間則將該產品列入推薦列表並予以推薦之。

為能推薦符合低忠誠度顧客之偏好產品以提升此特性顧客之推薦效能，推薦列表採用運用混合式推薦機制，有關低忠誠度顧客之推薦機制如圖 3。

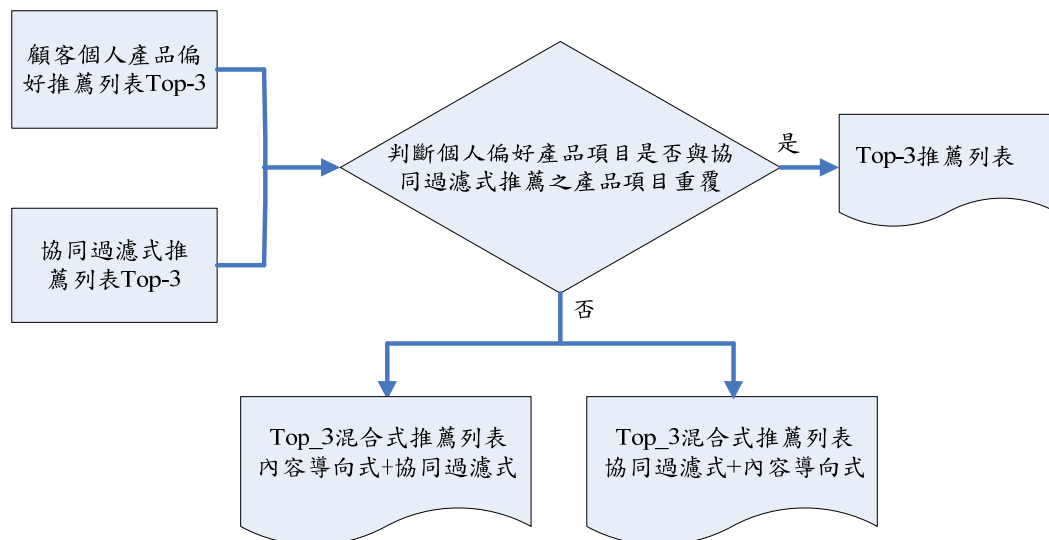


圖 3 低忠誠度顧客之推薦機制

經過顧客 RFM 分析階段後，被列入高忠誠度顧客群之顧客就是具有週期性之消費行為，因此針對此類特性之顧客，本研究依據其歷史交易紀錄，予以分析出個人購買產品之週期並求得各項產品之推薦區間，此推薦區間即為本研究針對高忠誠度顧客建立推薦列表之依據，並透過顧客最近一次購買產品類別 i 的消費時間，即可得知顧客下次對該類產品發生消費行為之時間，若有落入產品推薦區間則將該產品列入推薦列表並予以推薦之。詳細推薦流程如圖 4 所示。

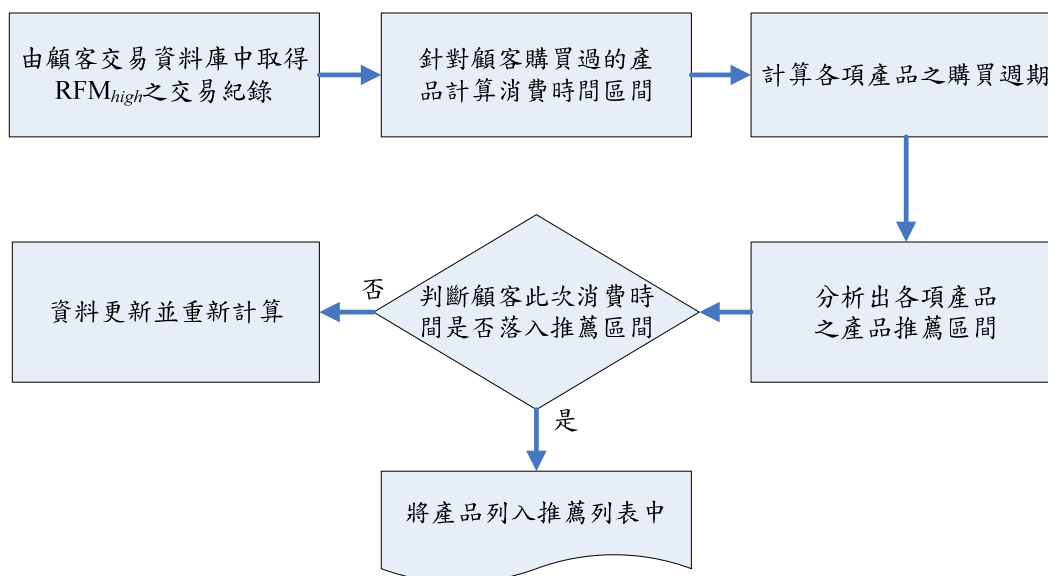


圖 4 高忠誠度顧客之推薦機制

肆、實驗分析

一、資料來源與工具

本研究資料來源為 Foodmart2000 食品超市銷售資料庫，實驗資料擷取交易資料表(TDT)、產品資料表(PDT)、產品類別資料表(PCDT)、顧客資料表(CDT)與時間資料表(DDT)作為分析及驗證研究方法之實驗依據。本實驗範圍取自 slaes_fact_1998 交易資料庫 1998/1/1 至 1998/12/31 之交易紀錄，共有 2,780 位顧客及 121,764 筆交易紀錄，並依據產品屬性歸類出 45 種類別(Category)。本研究利用 Microsoft office Access2003 作為資料之預處理與存取，並透過 Matlab 7.0 應用程式作為自組織映射圖網路分群之工具，最後再採用 Microsoft office Excel 2003 繪製實驗分析圖表並進行資料觀察與統整。

二、顧客 RFM 五等分分析

透過上一步驟之資料預處理後，可清楚地從顧客交易資訊中瞭解顧客過去何時購買哪些類別的產品，以及購買過的產品其數量與金額為何，藉由這些資訊與數據便可有助於本階段的顧客 RFM 分析。首先需取得每位顧客之原始 RFM 值，在得知三個指標的數值範圍後，即可利用公式(4)至公式(6)將三個指標予以正規化，其 RFM 五等分轉換結果之對照如表 2。

表 2 RFM 五等分轉換結果

value score	R-value(OR_u)	F-value(OF_u)	M-value(OM_u)
1	220 ~ 275	1 ~ 7	38 ~ 1253
2	165 ~ 220	7 ~ 15	1253 ~ 2506
3	110 ~ 165	15 ~ 22	2506 ~ 3759
4	55 ~ 110	22 ~ 30	3759 ~ 5013
5	1 ~ 55	30 ~ 41	5013 ~ 6303

將三個指標之數值做正規化處理轉換為 1~5 的分數後，即可加總三個指標之數值藉此得到每位顧客的 RFM 總分($Score_u$)，接著，利用公式(7)求得出所有顧客的平均 RFM 總分為 8，即 $\overline{Score} = 8$ ，並透過公式(8)之判斷式區分顧客之消費特性，若顧客 RFM 總分小於平均總分者視為低忠誠度顧客群(RFM_{low})，大於等於平均總分者則為高忠誠度顧客群(RFM_{high})。

三、分群結果分析

落入低忠誠度顧客群之顧客共有 1,440 位顧客，在本階段將低忠誠度顧客進行分群，以找出同群顧客之喜好排行榜並搭配顧客個人對產品之偏好進行混合式推薦。分群的輸入變數是將所有低忠誠度顧客之 RFM 值經五等分法轉換後，以每位顧客交易紀錄之 CR_u 、 CF_u 、 CM_u 值作為輸入變數，因此共輸入 1,440 筆，藉由自組織映射圖網路之參數設定即可得出對應之輸出單元，輸出單元即為網路之群聚結果，每一個輸出單元代表一個群組。

在進行自組織映射圖網路分群前，群數的數量(即拓撲之矩陣大小)必須在輸入變數前決定，為取得適當的群數數量，本研究共進行四個實驗，即 3x2、3x3、4x2 和 4x3 的群數設定，由群聚結果觀察出最佳群數，SOM 之分群實驗結果如表 3。

從四個實驗中顯示群數數量均介於 4 到 6 群，且有些群組內並無顧客存在，這代表本實驗之輸入變數具有明顯之特性，使得自組織映射圖網路(SOM)在進行分群時能清楚區別相似消費特性之顧客。本研究進一步分析群數為 5 與群數為 6 之分群結果可看出，其群組之人數相差懸殊，因此最後本研究採 4 群作為本實驗分析之依據。

表 3 自組織映射圖網路分群結果

Experiment Number	Ex 1	Ex 2	Ex 3	Ex 4
Matrix	3×2	3×3	4×2	4×3
C1	732	732	732	732
C2	0	0	0	0
C3	408	12	73	408
C4	0	0	195	0
C5	105	0	0	0
C6	195	195	408	0
C7	0	408	0	0
C8	0	73	32	0
C9	0	20	0	195
C10	0	0	0	20
C11	0	0	0	12
C12	0	0	0	73
Number of cluster	4	6	5	6

四、產品推薦區間結果分析

產品推薦區間分析結果，本研究以 Customer_id = 10250 其自 1998/1/1 至 1998/8/31 之交易紀錄為範例說明，其中擷取產品推薦區間分析所需之產品類別 (Product_category)、時間編號(Time_id)、購買數量(Unite_sales)等三項作為分析之依據。在顧客交易紀錄中，若產品類別之交易次數在兩次以下，無法求得交易時間區間與交易時間區間集合，因此本研究令交易次數門檻值 $\theta=3$ ，以下本研究以顧客編號 10250 購買第 4 類產品為範例說明產品推薦區間之分析步驟。

步驟 1：計算第 4 類產品之交易時間區間集合

利用公式(9)分別求出第 4 類產品之所有交易時間區間 DF_4^t ，並依序存入交易時間區間集合 DF_4 。

$$DF_4^1 = 888 - 735 = 153$$

$$DF_4^2 = 942 - 888 = 54$$

$$DF_4^3 = 966 - 942 = 24$$

$$DF_4 = \{153, 54, 24\}$$

步驟 2：求得第 4 類產品之購買週期

利用公式(10)至公式(12)計算顧客購買第 4 類產品之單個消費時間

ST_4^t ，再計算其最小購買週期 $ST \min_4$ 及平均購買週期 $\overline{ST_4}$ 。

$$ST_4 = \{51, 13.5, 4\}$$

$$\text{最小購買週期 } ST \min_4 = 4$$

$$\text{平均購買週期 } \overline{ST_4} = (51 + 13.5 + 4) / 3 = 22.8$$

步驟 3：分析產品推薦區間

首先利用 [公式\(13\)](#)與[公式\(14\)](#)求出該類產品之週期誤差 B_i 值與週期標準

差 σ_i ，接著藉由 [公式\(15\)](#)中的判斷式比較 B_i 與 σ_i 值之大小，以決定產品推薦區間之最大天數，並分析出該產品之推薦週期區間 $PRI_i[d_{\min}, d_{\max}]$ 。

$$B_4 = \overline{ST_4} - ST \min_4 = 22.8 - 4 = 18.8$$

$$\sigma_4 = 17.18$$

$$\text{由上述可得 } B_4 > \sigma_4, \text{ 故 } d_{\max} = ST \min_4 + \sigma_4 = 4 + 17.18 = 21.18$$

所以產品推薦區間為 $PRI_4[4, 21.18]$ 。

經過上述步驟求得第 4 類產品之產品推薦區間後，顧客 10250 購買過其他類之產品以此類推依序求得之產品推薦區間分析結果如 [表 4](#)。

[表 4](#) 產品推薦區間之分析結果

Product_category	$PRI_i[d_{\min}, d_{\max}]$
4	$PRI_4[4, 21.18]$
32	$PRI_{32}[8, 29.8]$
42	$PRI_{42}[2.7, 6.2]$
44	$PRI_{44}[3.3, 4.63]$
45	$PRI_{45}[0.7, 4.2]$

五、推薦列表結果

首先，針對低忠誠度顧客群之推薦列表項目，本研究結合自組織映射圖網路之協同式推薦，與顧客個人偏好之內容導向式合併推薦，利用此兩種推薦機制所分析出的結果，探討不同混合機制所產生的推薦項目對顧客推薦效能之情形。本研究設計兩種混合機制予以分析與評估。第一種是 CB+CF 機制，此種機制是取較多內容導向式所分析的項目，結合協同過濾式分析出的產品項目組成新的 Top_3 推薦列表。第二種混合機制是 CF+CB，此種機制則是取較多協同過濾式所分析的項目組成新的 Top_3 推薦列表。經由分析後，兩種混合式推薦機制所產生之混合式推薦列表內容列舉部分如 [表 5](#)。

表 5 低忠誠度顧客之混合式推薦列表

Customer_id	CB+CF Recommendation List	CF+CB Recommendation List
454	45, 36, 42	45, 42, 39
893	5, 13, 45	45, 42, 5
1765	45, 22, 42	45, 42, 22
2547	42, 4, 45	45, 42, 4
10104	45, 18, 42	45, 42, 18

本研究針對高忠誠度顧客群推薦列表的建立，是將顧客對各類產品所分析出的產品推薦區間作為推薦項目之依據，並可利用顧客最近一次購買產品類別 i 之消費時間為基準，以預測出顧客下次對該類產品所發生之消費行為時間為 $PRI[d_{\min} + NT_i, d_{\max} + LT_i]$ 。以表 4 之顧客 10250 其產品推薦區間分析結果為例，若顧客 10250 前一次消費時間為 1 月 29 日，可得知顧客 10250 對第 4 類產品之適時需求介於 2 月 4 日至 2 月 21 日，其他類產品則可以此類推分別求得產品之適時需求區間。之後，本研究將透過 1998/9/1~1998/12/31 之測試資料來驗證低忠誠度顧客與高忠誠度顧客其適性化推薦列表之推薦效能為何。

六、推薦效能評估

本研究針對低忠誠度顧客採用的推薦方式為混合式(Mix)推薦機制，並與傳統的內容導向式、協同過濾式比較其準確率、回應率與 F1 之推薦效能。除此之外，本研究實驗了兩種混合機制，即第一種為內容導向(CB)+協同過濾(CF)，與第二種協同過濾(CF)+內容導向(CB)，並將進一步探討兩種不同混合機制之推薦效能。實驗結果如圖 5 與圖 6。

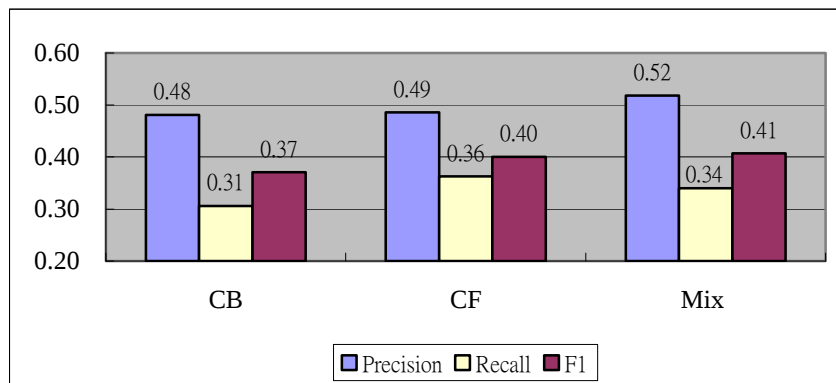


圖 5 比較 CB、CF 與混合式(Mix)推薦機制之推薦效能

由圖 5 之實驗結果所示，從 Precision 與 F1 兩項指標的數值，皆可看出混合式(Mix)推薦機制優於傳統的內容導向式與協同過濾式，然而在 Recall 值的表現上，

混合式推薦機制則低於協同過濾式推薦，產生此結果之原因為，由於 Recall 指標代表在顧客實際購買之產品中，命中系統推薦之產品所佔的比率，當分母皆相同的情況下，透過協同過濾式分析出的產品項目，其命中的個數多於混合式推薦，因此使得混合式推薦機制之 Recall 值產生較低的情況。

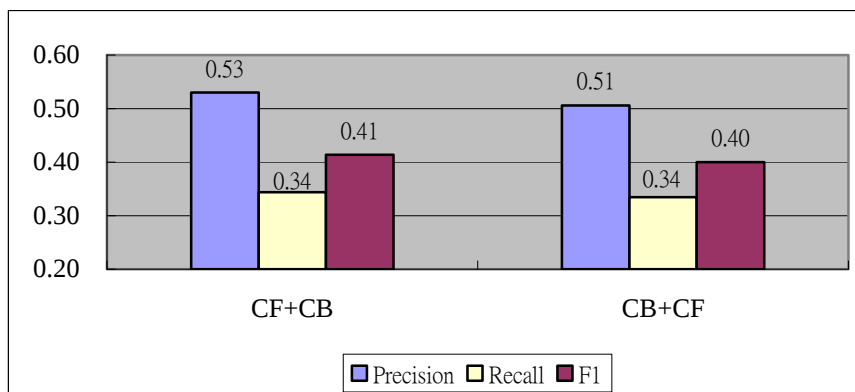


圖 6 比較不同混合機制之推薦效能

於圖 6 可發現，不論在 Precision、Recall 或是 F1 的表現上，CF+CB 之推薦效能皆優於 CB+CF，因為 CF+CB 之推薦機制所產生的推薦列表項目，取較多協同過濾式所分析的項目組合而成，而低忠誠度顧客購買的產品項目與數量較少，因此所分析出的個人偏好產品亦少，這代表著透過協同過濾式推薦給予顧客之排行榜產品，較能滿足顧客之潛在產品偏好，因此顧客實際購買的產品項目及個數，會購買較多由協同過濾式所分析出的產品，使得 CF+CB 能有較好的推薦效能。

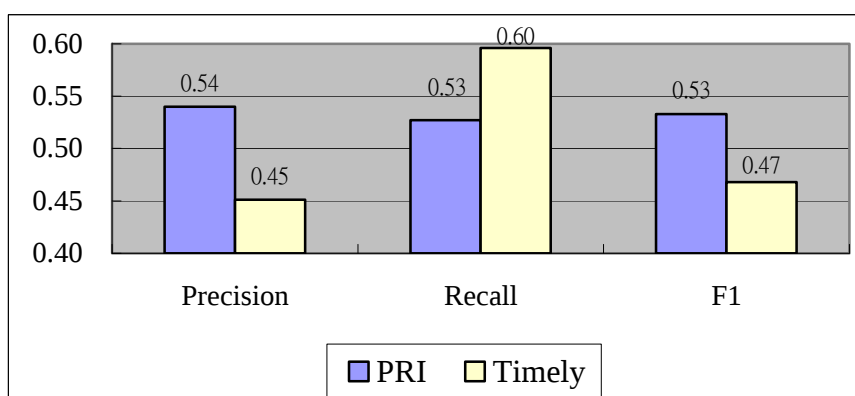


圖 7 產品推薦區間(PRI)與適時性(Timely)方法之推薦效能比較

圖 7 中顯示，從 Precision 與 F1 的結果可明顯看出，產品區間推薦機制確實能對顧客提升其推薦效能，表示產品區間推薦機制可推中顧客所需產品的比例高於適時性推薦方法。然而適時性方法在 Recall 指標上所呈現出的效能優於產品區間推薦機制，產生此情況之原因為，產品區間推薦機制設有交易總次數門檻值

($\theta = 3$)，未達此門檻值之產品將不會列入推薦列表中，當分母皆相同的情況下，此機制可產生的推薦項目較少，相對命中率便降低，因此使得適時性方法的 Recall 值高於產品區間推薦機制。

綜觀本研究適性化之實驗結果，針對低忠誠度顧客運用混合式推薦機制，其推薦效能與傳統內容導向式及協同過濾式相比，確實有較佳的推薦成效，代表本研究推薦之產品項目，能滿足顧客喜好並發掘顧客過去未曾購買過的潛在產品，藉此提升顧客忠誠度；另外，就高忠誠度顧客之推薦，從實驗結果顯示，產品推薦區間機制其整體推薦效能優於適時性推薦方法，表示本研究能確實改善適時性重覆推薦之缺失，因此更能準確掌握高忠誠度顧客之需求。

伍、結論

本研究考慮產品重覆購買以及產品潛在偏好的推薦，針對不同消費特性的顧客，提出一套基於 RFM 分析法之顧客適性化產品推薦機制。為避免推薦時機不適導致推薦效能不佳之缺失，重視顧客購買產品的潛在偏好，考慮顧客在不同時間對不同產品的週期需求，提供適時、適性的產品給予顧客，藉此提升顧客忠誠度並增加其再購意願。

在推薦效能評估中，實驗結果顯示，透過自組織映射圖網路分析出同群排行榜產品，並結合顧客個人產品偏好，強化低忠誠度顧客之推薦列表，可發掘出顧客未曾購買的潛在產品偏好，與傳統的內容導向式及協同過濾式推薦機制相比，確實能提升低忠誠度顧客之推薦效能，顯示混合式推薦機制相對於傳統推薦機制，能提供顧客較有效的產品推薦。另外，針對高忠誠度顧客群之推薦，本研究利用產品購買週期分析求得產品推薦區間作為推薦依據，其推薦效能在準確率與 F1 指標分析結果中可看出明顯優於過去學者之適時性推薦方法，代表產品區間推薦機制能適時地給予顧客推薦，且能確實掌握顧客對產品的購買需求，因此能達到更好的推薦成效。

陸、參考文獻

葉怡成，2003，*類神經網路模式應用與實作*，儒林圖書公司。

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), 61-70.

Ha, S. H., 2007. Applying knowledge engineering techniques to customer analysis in the service industry. *Advanced Engineering Informatics*, 21(3), 293-301.

- Hughes, A. M., 1994. Strategic database marketing. *Probus Publishing, Chicago*, 487-499.
- Kim, J. K., Cho, Y. H., Kim, W. J., Kim, J. R., and Suh, J. H., 2002. A personalized recommendation procedure for internet shopping support. *Electronic Commerce Research and Applications*, 1(3-4), 301-313.
- Kohonen, T., 1995. Self-organizing map, *Berlin:Springer*.
- Li, L. H., Lee F. M. and Liu, W. J., 2006. The timely product recommendation based on RFM method. *Proceedings of International Conference on Business and Information, Singapore, CD-E July 12-14*.
- Li, L. H., Hsu R. W., Hung S. S. and Tsai P. J., 2009. The personalized recommendation with bundling strategy based on product consuming period. *Proceedings of the 1th International Conference on Computational and Information Science*, 461-469.
- Liu, D. R. and Shih, Y. Y., 2005. Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information and Management*, 42(3), 387-400.
- Min, S. H. and Han, I., 2005. Detection of the customer time-variant pattern for improving recommender systems. *Expert Systems with Applications*, 28(2), 189-199.
- Oard, D. W. and Marchionini, G., 1996. A conceptual framework for text filtering. Technical Report EE-TR-96-25, CAR-TR-830, CLIS-TR-96-02, CS-TR-3643.
- Sarwar, B., Karypis, G., Konstan, J.h, and Riedl, J., 2000. Analysis of Recommendation Algorithms for E-commerce. *Proceedings of The 2nd ACM Conference on Electronic Commerce*, 158-167.
- Schafer, J. B., Konstan, J., and Riedl, J., 1999. Recommendation system in E-commerce. *Proceedings of The First ACM Conference on Electronic Commerce*, 158-166.
- Suh, E. H., Noh, K. C., and Suh, C. K., 1999. Customer list segmentation using the combined response model. *Expert Systems with Applications*, 17(2), 89-97.
- Yuan, S. T. and Tsao, Y. W., 2003. A recommendation mechanism for contextualized mobile advertising, *Expert Systems with Applications*, 24(4), 399-414.