

教育資料與圖書館學

Journal of Educational Media & Library Sciences

<http://joemls.tku.edu.tw>

Vol. 51 , no. 1 (Fall 2013) : 005-035

導入矩陣分群之視覺化圖書推薦系統

Visualized Book Recommender System

Using Matrix Clustering

郭俊桔* June-Jei Kuo*

Assistant Professor

E-mail: jjkuo@dragon.nchu.edu.tw

張瑞珊J ui-Shan Chang

Graduate Student

E-mail: changrs.tw@gmail.com

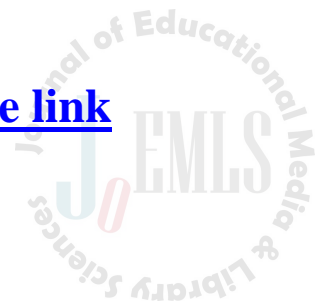
張育蓉 Yu-Jung Zhang

Graduate Student

E-mail: tatina8088vs@yahoo.com.tw

[English Abstract & Summary see link](#)

[at the end of this article](#)





導入矩陣分群之 視覺化圖書推薦系統

郭俊桔*

助理教授
國立中興大學圖書資訊學研究所
E-mail: jjkuo@dragon.nchu.edu.tw

張瑞珊

研究生
國立中興大學圖書資訊學研究所
E-mail: changrs.tw@gmail.com

張育蓉

研究生
國立中興大學圖書資訊學研究所
E-mail: tatina8088vs@yahoo.com.tw

摘要

傳統圖書推薦系統依據讀者過去的借閱紀錄，推薦相關書籍給讀者，也可藉由讀者所屬社群的資訊，推薦讀者從沒有借閱過的書籍。然而，讀者的閱讀興趣會隨著時間改變，借閱時間越近的圖書越能反應讀者當前興趣，每筆閱讀紀錄的重要性不可等同視之。圖書借閱紀錄高維度和稀疏的特性使得資料探勘的分群方法無法有效對應。再者，為使讀者可從推薦結果中有效地發現所需資訊，必須導入視覺化呈現技術。因此，本研究導入時間衰減因素，提出動態閥值矩陣分群，並導入主題地圖，以提高判斷圖書推薦適性之準確率。實驗結果證實視覺化圖書推薦系統比傳統圖書推薦系統具有更高滿意度，且雙層式主題地圖比單層式主題地圖更適合呈現推薦結果。

關鍵詞：時間衰減，矩陣分群，主題地圖，圖書推薦

*本文主要作者兼通訊作者。

緒 論

由於資訊爆炸和Web 2.0 (Maness, 2006)的興起，傳統圖書館漸漸轉型為數位圖書館。現代的人們已不再因資源不足或距離遙遠而產生資訊匱乏。相反的，由於網際網路的發達，絕大多數的人身處資訊洪流，看似唾手可得，但符合自己需求的資訊卻又總是淹沒在這些種類繁多的洪流中。面對大量的電子化資訊與資料，如何有效利用資訊系統幫助讀者過濾(filter)無用之資訊，在最短時間內取得所需圖書資源，已成為數位圖書館不斷努力的目標。因此，數位圖書館遭遇到前所未有的課題：如何有效管理大量館藏，提供讀者容易且快速資訊搜尋的環境(Peis, del Castillo, & Delgado-Lopez, 2008)。如果讀者利用以前使用過的關鍵字來搜尋新資訊，資訊檢索(information retrieval)或資訊過濾(Information filter)可有效提升讀者的資訊搜尋效率。然而，隨著電子商務個人化服務的研究與進展，推薦系統更被證實是一種解決上述課題有價值的工具(卜小蝶，1998；Ricci, Rokach, Shapira, & Kantor, 2011)。推薦系統(Adomavicius & Tuzhilin, 2005)是依據各種資訊，例如，購物履歷或商品屬性等，預測使用者的興趣後，推薦相關商品或資訊給使用者。有別於傳統的網站服務僅能提供資訊瀏覽服務的功能，讓讀者被動地接受資訊傳播，推薦系統更強調提供讀者回饋意見的互動和分享的管道。因此，在圖書資訊或電子商務領域，紛紛開發了各種推薦系統(Hwang & Lim, 2002; Aciar, Zhang, Simoff, & Debenham, 2007; Kuroiwa & Bhalla, 2007)。依據決定推薦過程的不同，可分成如下所示三種推薦系統：(一)協同式推薦系統(陳垂呈、陳幸暉，2011；Resnick, Lacovou, Suchak, Bergstrom, & Riedl, 1994)；(二)內容式推薦系統(Mooney & Roy, 2000)；(三)混合式推薦系統(Kumar, Ratnam, & Babu, 2010)。

在個人化服務的研究領域中，協同式推薦系統(collaborative filtering)更是最成功模型的一種(McLaughlin & Herlocker, 2004)。此系統的演算法是依據使用者的購物履歷，使用相似度計算由候選者中挑選具有購買履歷相似度較高之鄰居後，將這些人所購買但使用者尚未購買的商品推薦給使用者。然而，傳統的協同式推薦系統只有注意相似度運算，沒有考慮到隨著時間改變的興趣遷移(Chen, Jiang, & Zhao, 2010)而導致推薦結果常常偏離目前的興趣。除此之外，其他研究課題還有：(一)新讀者從來沒有借過書或無歷史借閱資料時，系統無法執行和其他讀者的相似度計算；(二)當一本新書還沒有累積到足以激發推薦的借閱數前，不會被推薦；(三)當圖書館整體書籍的借閱數太少或讀者數太少，即產生資料稀疏(sparsity)問題，浪費記憶體空間；(四)圖書館通常有大量的館藏和讀者，讀者一借閱紀錄呈現高資料維度(high dimension)，傳統的分群分析方法(Jain, Murty, & Flynn, 1999)無法有效執行讀者分群；(五)讀者基於個人的隱私，不願意提供借閱資料時，此方式無法有效運作。其中，(一)和(二)的問題一般被稱

為冷啟動 (cold start) 問題。另外，傳統推薦系統的研究大都只專注於提升推薦正確率相關演算法的改進，缺乏推薦結果的視覺化呈現，被推薦者無法針對推薦結果執行關連推導或呈現控制，以致於無法進一步了解推薦結果隱含的結構性，無法有效利用推薦結果 (Verbert, Parra, Brusilovsky, & Duval, 2013)。

因此，為了解決上述研究課題，本研究提出一個導入矩陣分群之視覺化圖書推薦系統。首先考慮時間衰減因素 (time decay factor)，動態調整圖書借閱次數以實際反應讀者興趣的遷移，接著使用動態閾值矩陣分群 (matrix clustering using dynamic threshold) 建立讀者社群。解決分群的運算效率課題後，導入主題地圖 (topic map) 呈現推薦結果，解決推薦結果呈現課題。為了驗證各個導入模組的功效，除了和傳統分群演算法比較分群成效外，並以問卷調查方式了解視覺化推薦結果呈現的讀者滿意度。本文結構如下：第一節說明研究背景、研究動機和研究目的；第二節則說明傳統矩陣分群及其問題點和解決方法，第三節提出使用動態矩陣的圖書推薦系統和說明其相關模組，第四節說明實驗和討論相關結果，最後第五節針對本研究的結論做綜合說明後，敘述未來可能的研究方向。

二、矩陣分群

分群分析 (clustering analysis) 在許多研究領域中是一種資料分析的基礎工具，主要是探討如何將資料予以分群 (cluster) 或分類 (classification)。其目的是針對一組以特徵向量呈現的資料，依其維度間的相似度將資料分成數個群體。其中每個群體內的元素具有高同質性，不同群集間的元素則具有高度異質性。因此，分群分析成為推薦系統不可或缺的工具。然而，分群分析方法種類繁多 (Jain et al., 1999)，主要可分成階層式 (hierarchical) 和分割式 (partitional)。階層式主要有完全鏈結 (complete link)，平均鏈結 (average link) 和單一鏈結 (single link) 等方法，其中單一鏈結和完全鏈結為較多人採行的方法。另一方面，分割式主要的方式則有 K-Means 演算法、ISODATA 演算法等，其中又以 K-Means 法最廣為使用。然而，階層式方式的分群結果會受其所使用的相似度計算方法影響且有不易解釋結果等課題；K-Means 方法則必須事先知道群集數、受初始群集中心的選擇影響很大，以及可能出現區域最佳解 (local maximum) 等課題。共同課題則是分群品質取決於資料量的大小。因此，如何解決上述傳統分群分析的問題，已成為開發協同式圖書推薦系統亟欲解決的課題。

(一) 矩陣分群的相關研究

網際網路上快速成長的電子商務，使得行銷策略漸漸轉向客製化，為了因應此趨勢，如何找出客人真正的需求已成為重要的研究課題。Oyanagi、Kubota 與 Nakase (2001) 提出矩陣分群法 (matrix clustering) 有效解決高維且稀疏矩陣

的網頁日誌 (web log) 分群。郭逸凡 (2003) 參考上述矩陣分群，提出類似的演算法來執行客戶關係管理 (Customer Relationship Management, CRM)，並得到令人滿意結果。CRM 主要將客人所有可能資料來源做一個綜合性整理與應用，針對每一位客人採取最佳化對應措施，以維持客人的忠誠度，不但使客人不會流失，更會幫企業帶來長期利益。由於圖書推薦系統和 CRM 有異曲同工之妙，皆是為客人或讀者推薦最適合的產品或圖書。再加上，由於電子商務中的商品－顧客矩陣大多是稀疏矩陣，矩陣分群可以有效處理稀疏矩陣的分群課題。因此，本研究導入矩陣分群來執行讀者的分群分析。

矩陣分群是由稀疏矩陣 (sparse matrix) 中抽出密實子矩陣 (dense submatrix) 來代表一個或數個群體，如圖 1 所示。如果把縱軸的 ABCDEFGH 想像成是八位讀者，而橫軸的 abcdefgh 想像成借閱圖書書名的話，右側的分群結果就可解釋為讀者 BFG 屬於具有相同閱讀喜好的社群。

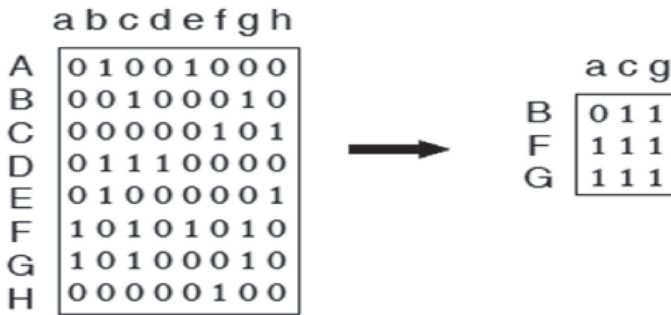


圖 1 高密實子矩陣的抽出 (Oyanagi et al., 2001)

(二) 矩陣分群的定義與演算法

矩陣分群使用如圖 1 所示的二值行列矩陣 (binary matrix) 為運算目標，行代表讀者，列代表借閱過之圖書；在行列矩陣 A 中， A_{ij} 表示讀者 i 與圖書 j 的借閱關係，若 $A_{ij}=1$ 表示讀者 i 借閱過圖書 j；反之， $A_{ij}=0$ 表示讀者 i 沒有借閱過圖書 j。因此，隨著讀者與圖書的規模變大時，矩陣內的值也會變得稀疏。觀察上述的二值矩陣，行列的讀者與圖書彼此相對的排列順序沒有特別的意義，所以行和列的排列順序可以彼此任意交換。

Oyanagi 等人 (2001) 為有效執行上述行列置換過程，提出了 Ping-Pong 演算法。Ping-Pong 演算法導入了資料探勘關連規則抽取的信賴度 (confidence) 和支持度 (support) 作為密實子矩陣搜尋問題的陣列評價目標。執行矩陣分群時，支持度被定義成子矩陣的值为 1 的元素個數；信賴度則被定義成子矩陣的密度。換句話說，矩陣分群的目的是求得面積或密度最大化的分群結果。因此，其搜尋問題可分為如下所示兩種評價函數 (evaluation function)：

1. 密度最大化：在指定面積上，搜尋密度最大的子矩陣。
 2. 面積最大化：在指定密度上，搜尋面積最大的子矩陣。
- Ping-Pong 演算法主要結構如下所示。

```

while (convergence) {
    row_to_col ();
    prune_col ();
    col_to_row ();
    prune_row ();
}

```

此演算法的輸入是二值矩陣，輸出則為密實矩陣；反覆執行密實子矩陣的抽取，一直到行和列都不會再變化為止。每次循環的步驟是首先，執行 row_to_col() 函數由行 (row) 激發列 (column)，並計算每列的節點數 (count) 後，執行 prune_col() 函數，刪除節點數小於等於刪除閾值的列。接著，執行 col_to_row() 函數由列激發行，並計算每行的節點數後，執行 prune_row() 函數，刪除節點數小於等於刪除閾值 (threshold) 的行。

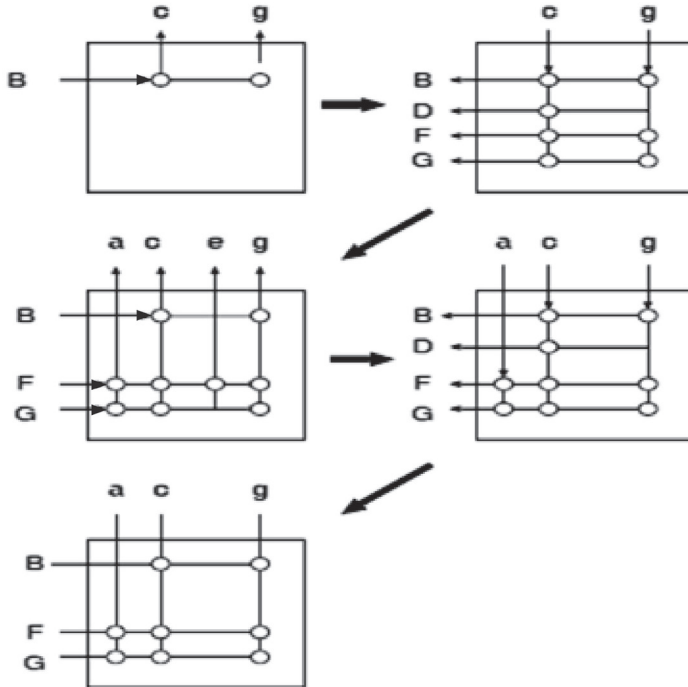


圖2 Ping-Pong 演算法 (Oyanagi et al., 2001)

以圖1的二值稀疏矩陣為例，其Ping-Pong演算法執行示意圖如圖2所示。首先，設定刪除閾值為1並由行B開始激發 (activate) 後，連結到列c和g，計

算每列上的節點數。(row_to_col步驟)然後，執行列c、g的刪除，此時雖然各列的count值為1，由於剛啟動而不執行列刪除。(prune_col步驟)接著，執行列c、g的激發。列c連結了行BDFG和列g連結了行BFG後，計算每行上的count值。(col_to_row步驟)由於行D的count值為1(等於刪除閾值)其他為2，因而執行刪除行D(prune_row步驟)。接著，激發行BFG，並計算列aceg的count值。(col_to_row步驟)然後，由於列e的count值為1，執行列刪除列e。(prune_row步驟)接著，激發行acg。(row_to_col步驟)此時，由於行和列都不再變化，所得到密實子矩陣是由行BFG和列acg所構成。接著，計算此密實子矩陣的1之個數和密度，分別為8和 $8/9=0.89$ 。依據選定密度和面積的評價目標，即可判定此擷取子矩陣是否為密實子矩陣。

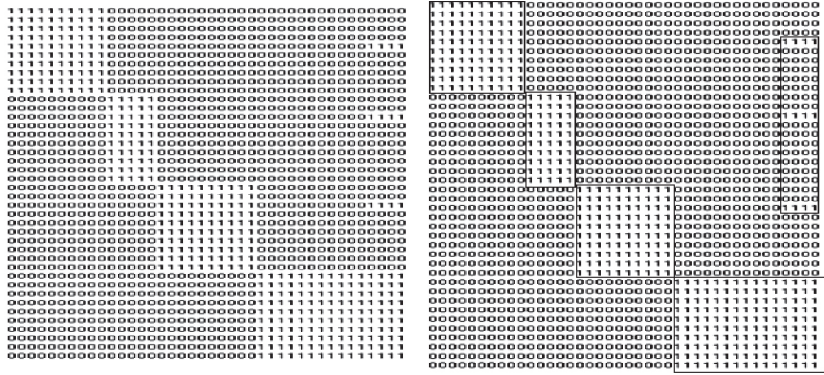
參考上述Ping-Pong演算法，Kuo與Zhang(2012)為解決Ping-Pong演算法只能得到一個而不是所有的密實矩陣，首先提出靜態閾值矩陣分群演算法。然而，其實驗結果發現固定密度閾值造成群體被強制分割，產生錯誤分群。接著，分析實驗結果發現，當無法在繼續產生密實矩陣時，如果同時導入遞減的密度閾值和刪除閾值，可避免由於固定閾值所產生之強制分割的錯誤。換句話說，由於密度閾值或刪減閾值的限制造成剩餘矩陣元素無法再合成任何密實子矩陣時，遞減閾值而使得這些矩陣元素可以再次被合成密實子矩陣。因此，Kuo與Zhang(2012)修改其靜態閾值矩陣分群演算法為如下所示的動態閾值分群演算法。

動態閾值矩陣分群演算法：

- (1)輸入二值的稀疏矩陣A和設定相關參數和閾值，例如密度值等
- (2)使用亂數決定尚未激發的入口行數
- (3)執行Ping-Pong演算法
- (4)由矩陣A刪除已擷取的密實子矩陣資料
- (5)如果已無法抽出密實矩陣或矩陣結束已為空矩陣，執行(6)；否則，執行步驟(2)
- (6)如果稀疏矩陣A不是空(NULL)的，遞減密度參數和激發值
- (7)如果密度參數或激發值大於其預定臨界值，執行步驟(2)
- (8)輸出擷取的所有密實子矩陣
- (9)結束

本研究為了評估動態閾值矩陣分群的效果，使用如圖3(a)所示的二值稀疏矩陣作為輸入。接著，為了找出矩陣分群中各種參數的最佳值，分別使用不同的值域和遞減值的組合探討參數之最佳值：密度閾值(值域：0.1~1.0，遞減值：0.1)，刪減閾值之臨界值(值域：1~10，遞減值：1)，密度閾值之遞減值(值域：0~1，遞減值：0.1)和激發值之遞減值(值域：0~10，遞減值：0.5)。執行動態閾值矩陣分群後，得到如圖3(b)所示之正確分群結果。最好的密度閾值和

刪減閾值之臨界值分別為0.6和4。另一方面，密度閾值和激發值遞減值則分別為0.1和1。因此，在下一節的視覺化圖書推薦系統是導入上述最佳參數組合之動態閾值矩陣分群演算法於分群運算。

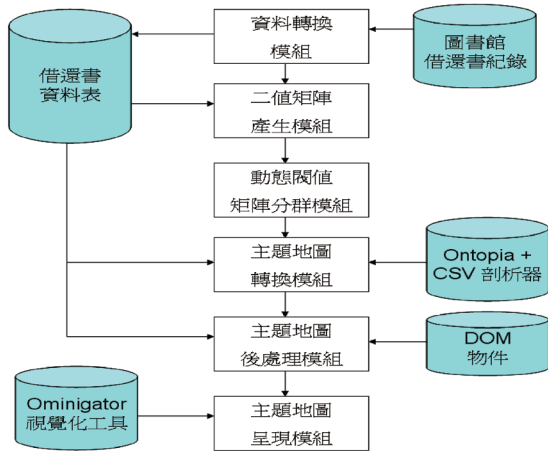


(a)輸入二值矩陣 (b)輸出密實矩陣

圖3 使用動態閾值演算法之密實子矩陣

三、基於動態閾值矩陣分群之視覺化圖書推薦系統

圖4為本研究之圖書推薦系統架構，包括資料轉換模組、二值矩陣產生模組、動態閾值矩陣分群模組、主題地圖轉換模組、主題地圖後處理模組和主題地圖呈現模組等六個模組。其處理流程為資料處理模組將圖書館借還書紀錄讀入後，選擇需要的屬性並擷取相關的資料後，儲存成為借還書資料表；接著，二值矩陣產生模組導入時間衰減因素，調整借閱書籍次數後，參考閾值將借還書資料表轉換成二值矩陣後，動態分群模組依據二值矩陣內容處理並輸出CSV



(彩圖請見電子檔)
圖4 基於動態閾值矩陣分群之圖書推薦系統架構圖

格式的分群結果；然後，主題地圖轉換模組使用 CSV 剖析器由 CSV 格式之分群結果抽出必要的資料，並使用 Ontopia 工具將其轉換成 XML 格式之主題地圖；接著，主題地圖後處理模組使用文件物件模型 (Document Object Model, DOM) 模組修正主題地圖的同一物件參照問題後，主題地圖呈現模組使用 Ontopia (Pepper, 2005) 的 Omnigator (2009) 內部套件顯示主題地圖以及讓讀者可以自由的展開或縮小主題地圖。以下各節將進一步敘述各模組的處理和所需的相關資料。

(一) 資料轉換模組

圖書館的借還書紀錄通常包含很多欄位，然而並不是所有欄位都是圖書推薦系統所需。此模組將刪除不需要的欄位，儲存剩餘的欄位資料。接著，將同一個讀者的所有相關紀錄集中並合併相同一本書的紀錄 (例如刪除續借資料，只保留第一次借書和最後一次還書資料)，分別針對每一本書計算其借閱次數、開始借閱日期、最後還書日期和借閱期間，並儲存於如圖 5 所示的借還書資料表中。然後，使用 MARC_ID 從書籍紀錄中，擷取書名、作者、出版社和出版年份，也儲存於借還書資料表中。

PROCESS_USER_ID	ACTIVITYRNO	ACC_NO	CLASS_N	MARC_ID	MEMO	LAN borrow_count	start_date	return_date	dura			
2008/10/6	CIR1	借	79414001	515992	496	326898	應還日:2008/11/05	chi	1	2008/10/6	2008/10/31	26
2008/10/6	CIR1	借	79414001	886145	24.6	284816	應還日:2008/11/05	chi	2	2008/10/6	2008/11/4	30
2008/10/6	CIR1	借	79414001	927678	24.6	135835	應還日:2008/11/05	chi	1	2008/10/6	2008/10/6	1
2008/10/6	CIR1	借	79414001	929818	24.6	360318	應還日:2008/11/05	chi	2	2008/10/6	2008/11/4	30
2008/10/6	CIR1	借	79414001	929819	24.6	360318	應還日:2008/11/05	chi	2	2008/10/6	2008/11/4	30
2008/6/1	CIR2	借	79414001	532168	23.68	360310	應還日:2008/09/15	chi	2	2008/6/1	2008/10/6	128
2008/6/1	CIR1	借	79414001	553757	494.2	524987	應還日:2008/09/15	chi	1	2008/6/1	2008/7/9	39
2008/6/1	CIR1	借	79414001	619663	541.7	8891	應還日:2008/09/15	chi	1	2008/6/1	2008/7/9	39
2008/6/1	CIR1	借	79414001	643717	170	14240	應還日:2008/09/15	chi	1	2008/6/1	2008/7/9	39
2008/6/1	CIR2	借	79414001	674876	24.5232	77383	應還日:2008/09/15	chi	2	2008/6/1	2008/10/6	128
2008/6/1	CIR1	借	79414001	856054	170	80093	應還日:2008/09/15	chi	1	2008/6/1	2008/7/9	39

title	author	publisher	publish year
圖書資訊利用教育:國小階段	林菁	五南	2001
圖書館的利用,國民小學篇	曾雲娥	國家圖書館	1999
中小學圖書館利用教育手冊(曾雲娥	台北市政府教育局	
中小學圖書館利用教育手冊(曾雲娥	台北市政府教育局	
閱讀活動策劃與行銷研習班	中華民國圖書館	中國圖書館學會	9999
組織行為學	羅賓斯 (Robbins,	華泰文化	2006
社會心理學	西爾思 (Sears, D	五南	1989
認知心理學	鍾聖校	心理	1990
臺灣地區的兒童圖書館	鄭雲坎	漢美	1993
認知心理學:理論與應用=Cog	鄭麗玉	五南	1993

圖5 借還書資料表

(二) 二值矩陣產生模組

為了降低二值矩陣的維度，本研究不使用原來的圖書分類號，只參照「中國圖書分類表」(賴永祥，2001)將借還書資料表中的圖書分為十大類。因此，參照借還書資料表，依序將每個讀者所借的圖書，彙整各類圖書借閱的總數和總借閱次數。除了上述彙總資料外，依據分析需要由相關資料表抽出資料。例

如，使用借閱證號碼由讀者資料表抽出讀者的姓名、性別、出生年份或在職與否等個人資訊。

圖書被借閱的時間越接近，表示對讀者而言是越重要或越感興趣的圖書。反之，表示越不重要或越不感興趣。這種因素可稱之為時間衰減因素。傳統圖書推薦系統都使用線性函數的時間衰減因素 (Gong & Cheng, 2008; Zhang & Ye, 2012)。另一方面，Chen 等人 (2010) 發現人類遺忘興趣的速度遠比線性快，提出基於時間權重函數 (time-based weighting function) 和使用參數 m 的大小來調整遺忘速度，並有不錯的實驗結果。為探討這個課題和了解遺忘速度的影響，參考 Chen 等人 (2010) 的最佳函數值，提出如公式(1)所示的指數函數作為時間衰退因素，在計算每一類的總借閱次數時，使用時間衰退公式計算之。TD-SUM 為某一欄位的彙總，TD $_i$ 則為距今年 i 年前的值。Decay_Factor 則是衰退率，其值域為 0~1 之間，例如第一年的話，由於是 0 次方，所以使用原來的值。接著，對每一個資料欄位訂定一個閾值後，將值轉換為 0 或 1。

$$TD-SUM = \sum_{i=1}^n (decay_factor)^{i-1} \times TD_i \quad (1)$$

本研究使用不同的閾值 (0.99~0.90)，分別製作測試者的測試表單 (推薦書單) 後，由測試者的滿意度發現最佳的衰退率數值為 0.95。

(三) 動態閾值矩陣分群模組

本模組為有效執行使用者分群，導入上述之動態閾值矩陣分群演算法 (Kuo & Zhang, 2012)。矩陣分群的問題可被視為共分群 (co-cluster) 問題 (Angiulli, Cesario, & Pizzuti, 2006)，被定義為：給予一個 $I \times J$ 大小的二值稀疏矩陣 $A(I, J)$ ，找出最多滿足同質性特性 (homogeneity characteristics) 的長方密實形矩陣。其評價函數 Q 如公式(2)~(4)所示： $a_{i,j}$ 為第 i 行的平均值， $a_{i,j}$ 為第 j 列的平均值， $v_{i,j}$ 是稀疏二值矩陣中 1 的總數， $Q(I, J)$ 使用每行和每列的 1 的平均值，偵測最大且密實的子矩陣。再者，如果只使用一次方的話，將無法區別有無共享行或列的子矩陣。以圖 6 為例，如果只使用單次方的話，行或列的總和都是 4；然而左邊較密實的圖 (a)，如果使用平方的話，(a) 和 (b) 的總和分別為 16 和 12。如果想要放大這種現象，可使用三或以上的次方。本研究的實驗結果仍以兩次方為最佳，評估函數如公式(4)。

1	1	0	0	1	0	0	1
1	1	0	0	0	1	1	0

(a)

(b)

圖 6 具有相同行或列平均值的不同矩陣 (Angiulli et al., 2006)

$$a_{ij} = \frac{1}{|J|} \sum_{j \in J} a_{ij}, \text{ and } a_{ij} = \frac{1}{|I|} \sum_{i \in I} a_{ij} \quad (2)$$

$$v_{IJ} = \sum_{i \in I, j \in J} a_{ij} \quad (3)$$

$$Q(I, J) = \frac{\sum_{i \in I} (a_{ij})^2 + \sum_{j \in J} (a_{ij})^2}{|I| + |J|} \times v_{IJ} \quad (4)$$

由於分群的問題屬於NP的問題，所以必須導入適當的啟發函數 (heuristic function) 於上述分群問題。本研究使用的啟發函數不考慮上述個數和密度的限制，把尚未分群的 I 的數值視為同一群並使用公式(4)來計算啟發函數值 Q'。評價函數如公式(5)所示，其意義為把 j-1 個已分群好的 Q(I, J) 值和相關啟發函數 Q' 的和。其中，i 值指的是目前已組合完成的群組數，j 是正在執行第 j 個群組合併。因此，目標函數為 $\arg \max_{n \in I} Q''(n)$ 。

$$Q''(j) = \sum_{i=1}^{j+1} Q_i(I, J) + Q'(j) \quad (5)$$

換句話說，使用目標函數和貪婪搜尋方式 (Russell & Norvig, 2003)，執行上述導入動態閾值之矩陣分群演算法，可求得具有最高評價函數的動態分群結果。

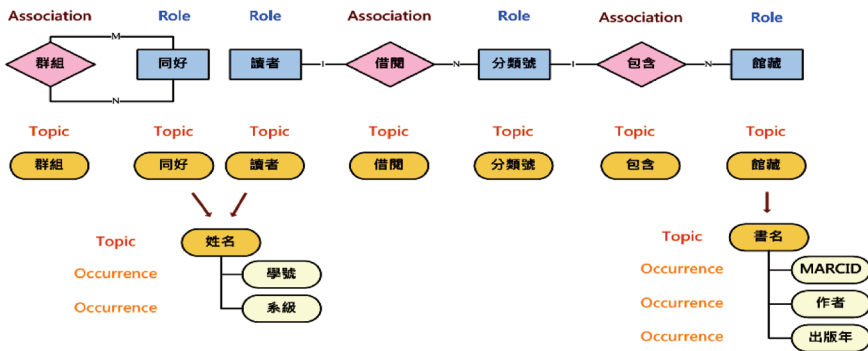
(四) 主題地圖轉換模組

主題地圖 (TopicMaps.Org, 2005) 是一種被應用在建置網站資源中心的知識管理整合技術，適合用於知識管理，呈現更多的資訊內容給使用者，如關係推導 (relation propagation)。讀者可以藉由了解自己的所屬群組，不需依靠其他圖書查詢系統就能發現新的可能需求書籍，達到圖書推薦的目的。因此，「主題地圖」相關視覺化技術加上智慧型人機介面，應可解決圖書推薦系統的相關課題。

要將資料自動轉換成主題地圖，首先要將動態閾值矩陣分群模組輸出的群組資料輸入本模組。接著利用 CSVParser (2011) 套件實作剖析群組資料和借還書資料表 (CSV 檔) 後讀取需要的資料內容，以 Ontopia (Pepper, 2005) 提供的 Topic Maps Engine 和 API 建立主題地圖。最後，將初步完成的主題地圖輸出成 XTM 格式以便使用網頁瀏覽。主題地圖是由主題 (topic)、關聯 (association) 和資源指引 (occurrence) 三大部分組合而成。使用者可從主題名稱辨識內容，並從資源指引獲得更詳細的資訊。主題表達單一概念，通常僅儲存文字內容。資

源指引可用來描述主題，除儲存文字資料外，還可儲存超連結資訊，向外指引資源所在位置。關聯名稱是以主題形式儲存，利用自身連結兩個不同主題之間的關係。

本研究使用的讀者和館藏兩種主題需要使用資源指引描述細節資訊。前者意指學生身分的讀者，後者為館藏圖書。讀者主題原本只顯示學生姓名，添加資源指引後可另外顯示學號和系級資訊。館藏主題原本則只顯示題名，加入資源指引後可顯示出MARC_ID、作者和出版年等書籍資訊。提供讀者和館藏的細節資訊有助於使用者辨識主題內容，避免誤認相同主題名稱但內容資訊不同的物件。主題地圖組成元素如圖7所示：銘黃色橢圓形為主題，描述物件概念；淺黃色橢圓型為資源指引，也就是附屬於主題之下的屬性資料；淺藍色長方形為關聯內的角色 (role)，也是一種主題；淺紅色菱形為關聯，描述兩個主題之間的關係。



(彩圖請見電子檔)

圖7 圖書推薦之主題地圖記號

(五) 主題地圖後處理模組

主題地圖轉換模組產出的主題地圖在呈現上有一個重大瑕疵：相同內容的主題資料會重複出現在同一張主題地圖。Ontopia套件內的預設函式在建立主題和關聯時並未比對物件內容，僅以TopicID是否相同的方式辨識重複物件，使得承載相同內容但於不同時間點產出的主題會以不同的TopicID標記，被認為是不同的物件，造成空間的浪費和主題間關係的斷裂。導致同一讀者的資料可能會重複儲存於多個物件中，或某一關聯欲指向特定主題物件時卻無法參照到真正想連結呈現的主題物件。

為解決上述問題，以DOM套件處理主題地圖檔，分批去除重複物件。首先，讀取初始資料建立物件TopicID對照表。接著對主題地圖進行檢測，搜尋比對重複的主題和關聯物件，將錯誤參照的物件予以修正和刪除。根據主題類別的不同，分成以下模組執行去除重複的動作：1. 去除分類號主題重複，2. 去

除讀者主題重複，3. 去除群組主題重複，4. 去除館藏主題。當所有可能產生重複的主題都刪除並檢測完畢後，輸出更新後的主題地圖供後續研究呈現之用。

(六) 主題地圖呈現模組

Omnigator 工具 (2009) 是 Ontopia 內部套件 (Pepper, 2005) 之一，可用於匯入/匯出、瀏覽、編輯及管理主題地圖。本研究除了使用 Ontopia 的 Topic Maps Engine 和 API 製作主題地圖外，還利用 Omnigator 能以視覺化方式呈現主題地圖的功能，讓圖書推薦的呈現更具直覺性，且便於使用者瀏覽與評估實驗結果。

由於本研究探討主題為大學圖書館的圖書推薦系統，主題地圖主要使用三種關聯，分別是借閱、包含及群組。群組各自以數字命名以示區別，關聯主題則以群組為首字命名。關聯內角色包含分類號、同好、書名和讀者。角色名稱與關聯主題相關，例如借閱關聯需要由讀者和分類號二種角色才能表示出「讀者借閱某個分類的書籍」的意涵，而包含關聯就需要分類號和書名表示「每本書歸屬於某個分類之下」的意義。群組的關聯較特殊，具有相同喜好的讀者才會被聚集在一起，沒有上下位階層的關係，可直接以「讀者 A 和讀者 B 是同好」的方式表示。最後，資源指引類型的部分分別是標示讀者資訊的學號、系級，以及標示館藏書目資料的 MARC_ID、作者、出版年和分類號等六種。這些屬性資料都是為了彌補主題名稱所包含資訊可能不足以識別的問題，因此可隨實際需求增刪修改，使用上十分自由。如圖 8 所示右側為由館藏和分類號兩種主題組成的關聯，左側視窗內是該書的主題名稱和詳細書目資料的主題都可在主題地圖中呈現。

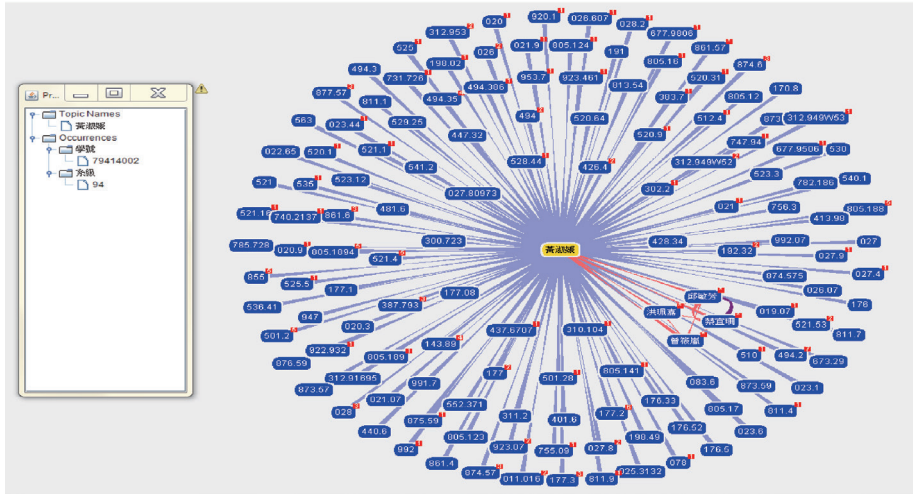


(彩圖請見電子檔)

圖 8 主題地圖之屬性呈現

單一讀者與其借閱過之分類號的關聯特性為借閱過的資料越多，讀者主題週遭的節點越多。如圖 9 所示，右下方桃紅色線連結與中心位置讀者相同群組的其他讀者。若點選此主題，主題地圖會自動展開相關主題。不同類型的關聯

會以不同顏色標示，方便讀者閱讀與辨識。



(彩圖請見電子檔)

圖9 主題地圖之相關主題呈現

四、實驗結果

為了方便評估上述導入動態閾值矩陣分群和主題地圖之圖書推薦系統的效能，除了實施讀者滿意度問卷調查外，還進一步經由訪談了解讀者的使用經驗。本研究以中興大學圖書資訊學研究所94~99級的研究生為研究對象。學生數總計71名，扣除借書還書太少者，選取59名同學之借還書資料為實驗語料。

(一) 系統開發環境

本實驗在作業系統使用Windows XP或Windows 7下開發圖書推薦系統的所有相關模組。矩陣分群相關的程式，使用C語言開發和使用Excel製作與維護如圖5所示借還書資料表。主題地圖的相關程式，則以J2SE JDK配合Eclipse進行程式執行與除錯；另以Ontopia-5.1.3配合Apache Tomcat進行主題地圖的建立和呈現；利用JAVA的CSVparser套件讀取及剖析CSV檔案，擷取需要部分的資料存入主題，並使用W3C DOM文件物件模型進行主題地圖的增刪修改，去除重複主題與修正關聯參照。

(二) 資料集

中興大學圖書館提供2003/1~2010/11的讀者借還書紀錄，共539,598筆，其中每一筆紀錄是由14個欄位所構成，例如事件日期、事件時間、操作者ID、交易代碼(借、還、續借或預約)、借書證代碼、登錄號、分類號、館別、MARC_ID、MARC_TYPE、應還書日/還書日、特藏類型和資料語言等14個欄位。為了不產生稀疏二值矩陣，本研究之分類號都只採用第一個數字。另外，還提供讀

者資料表(78,770讀者資料)、圖書MARC資料(約3,073本書)、MARC_ID與書名對照表等以方便製作借還書資料表。

首先,由圖書館借還書紀錄擷取圖資所同學的借還書紀錄,共有14,762筆。經過資料轉換模組處理後,讀者資料表有4,908筆紀錄。讀者資料表的欄位有15個欄位,如性別、出生年次、是否在職生、0類(總類)書籍數、1類(哲學類)書籍數、2類(宗教類)書籍數、3類(科學類)書籍數、4類(應用科學類)書籍數、5類(社會科學類)書籍數、6類(史地類)書籍數、7類(世界史地類)書籍數、8類(語言文學類)書籍數、9類(藝術類)書籍數、借閱書籍總數、距最後一次借閱經過幾天等。這些屬性都有協助分群的功用,例如在校生由於修課的關係,常會借閱相同分類的圖書,所以使用「距最後一次借閱經過幾天」的屬性值可用來區分該研究生是否仍是在學生等。其中,每個類別的書籍數,分別計算原始總數和導入時間衰退因素的總數,以執行不同的實驗。另外,為將資料集轉換成為二值矩陣,本研究測試各種閾值組合,得到如下所示最佳閾值條件。在二值矩陣產生模組中,性別:男(1)和女(0);出生年次:1981年(含)後為(1)和1981年前為(0);在職與否:一般生為(1)和在職生為(0);每類書籍數 ≥ 10 為(1), <10 為(0);借閱書籍總數 ≥ 50 為(1), <50 為(0);距最後一次借閱經過天數 ≤ 700 為(1), >700 為(0)。括號中數值為符合前述條件之對應的二值數值,例如,若受測者為男生,其性別屬性值為1,或者受測者借閱0類書籍數超過10本的話,0類書籍數屬性值為1。因此,得到 59×15^1 的二值矩陣後,作為矩陣分群模組的輸入矩陣。

(三) 推薦書籍的滿意度

使用上述 59×15 的二值矩陣作為輸入矩陣,執行動態閾值矩陣分群演算法後,最佳分群數為15。為了比較推薦圖書系統的效能,使用K-Means和Complete Link分群法(Jain et al., 1999)實作兩個分群系統,與本研究建議的矩陣分群方法進行比較。其中,K-Means法的群數K則設定為和矩陣分群之最佳群數相同($K=15$)。另一方面,由在校生中隨機抽取四位來進行滿意度調查,其中兩位為研究所三年級,另兩位為研究所二年級。接著,參照這四位同學的分群結果,分別製作如圖10所示的測試表單,表單上列出所有同一分群同學借閱過的書籍(去除重複和自己已借過書籍),每筆紀錄的欄位內容為書名、出版社、出版年份和語言外,列出兩個選項讓同學勾選。「有興趣」表示對此書有需求,馬上借;「Don't know」表示有興趣,以後才會借。相反的,完全沒有興趣則不要勾選,兩者只能擇一勾選。

1 59為研究对象人數,15為性別、出生年次、是否在職生、0類書籍數、1類書籍數、2類書籍數、3類書籍數、4類書籍數、5類書籍數、6類書籍數、7類書籍數、8類書籍數、9類書籍數、借閱書籍總數,和距最後一次借閱經過幾天等15個屬性值。

1	author	publisher	year	lang.	有興趣	Don't know
2	實踐大學圖書館/2000年讀者服務研討會	實踐大學圖書館	1999	chi		
3	Brown, William Houston./Ahern, Lawrence R.	Thomson/West,	2006	eng		
4	國家圖書館圖書採選編目部/全國文獻編目工作研討會論文集	北京圖書館	2006	chi		
5	胡歐蘭教授七秩榮慶籌備小組	文華圖書館管理	2005	chi	✓	
6		開明出版	2003	chi	✓	
7	施啓元 (應用美術)	上智科技	2006	chi		
8	董大偉	文獻資訊	2003	chi		✓
9	Rasmussen, Robert K.	Foundaton Press,	2007	eng		
10	Powell, Ronald R./Conaway, Lynn Silipigni.	Libraries Unlimited,	2004	eng		
11	McKinley, Robin.	Harper Trophy	1993	eng		
12	Strickland, Dorothy S./Alvermann, Donna E.	Teachers College Press,	2004	eng		
13	Scarberry, Mark S.	Thomson/West,	2006	eng		
14	Francis, William H./Francis, William H.	Thomson/West,	2007	eng		
15	Richmond, Velma Bourgeois.	McFarland & Co.,	2004	eng		
16	Van Tassel-Baska, Joyce/Stambaugh, Tamra.	Pearson/Allyn and Bacon,	2006	eng		
17	國家實驗研究院科技政策研究與資訊中心	國家實驗研究院科技政策研究與資訊中心	2006	chi	✓	

(彩圖請見電子檔)

圖10 測試表單

上述四位測試同學經過 15 分鐘簡短說明後，每位收到六份測試單（2 種時間衰減權重 × 3 種分群方法）並開始勾選。然而，測試學生並不知其所測的組合種類。其中，使用和不使用時間衰減組別分別表示使用和不使用公式 (1)，調整各類別圖書的總數。另一方面，計算正確率時，「嚴格」：有興趣，「寬鬆」：有興趣 + 不知道。測試結果以選取平均值表示。假設推薦書籍有 100 本，若在「有興趣」勾了 50 本，在「Don't know」勾了 15 本的話，嚴格的分數為 0.5，而寬鬆的分數為 0.65。表 1 為各測試者的測試結果（正確率），表 2 則表示結果的平均值。正確率越高，表示推薦的圖書越是符合使用者的資訊需求。

表 1 圖書推薦測試結果

測試者	時間 衰退	方法	種類		
			K-Means Method	Complete Link	Proposed Method
Tester1	不使用	嚴格	0.1934	0.1757	0.6415
		寬鬆	0.2378	0.2204	0.6843
	使用	嚴格	0.2268	0.1173	0.7153
		寬鬆	0.2823	0.1404	0.7495
Tester2	不使用	嚴格	0.3644	0.1434	0.1916
		寬鬆	0.4745	0.2565	0.3770
	使用	嚴格	0.5454	0.4935	0.2251
		寬鬆	0.6212	0.5714	0.4188
Tester3	不使用	嚴格	0.4012	0.3503	0.3971
		寬鬆	0.4349	0.3875	0.9266
	使用	嚴格	0.3900	0.4408	0.5000
		寬鬆	0.4087	0.4534	0.9716
Tester4	不使用	嚴格	0.2084	0.4329	0.5027
		寬鬆	0.2568	0.4639	0.5297
	使用	嚴格	0.1778	0.4385	0.6242
		寬鬆	0.2147	0.4705	0.7100

表2 實驗結果平均值

時間 衰退	方法	種類		
		K-Means Method	Complete Link	Proposed Method
不使用	嚴格	0.2918	0.2755	0.3075
	寬鬆	0.3510	0.3320	0.6294
使用	嚴格	0.3350	0.3725	0.5161
	寬鬆	0.3817	0.4089	0.7124

參照表1和表2所示的實驗結果，本研究的圖書推薦方法顯然比K-Means或Complete Link分群法更能推薦與讀者興趣相近的圖書。換句話說，針對高維度且稀疏的二維矩陣，矩陣分群的分群效果優於傳統的分群演算法。此外，由於使用時間衰退組的平均值都大於不使用時間衰退組者，更可證明導入時間衰退因素是必要和有效的。另一個有趣的發現是傳統的K-Means和Complete Link在不導入時間衰退因素時，K-Means優於Complete Link。相反地，如果導入時間衰退因素時，則得到完全不同的結果。因此，其可能的原因是導入時間衰減因素，可使推薦圖書的種類更趨近使用者的興趣，即使使用Complete Link也可以得好的分群效果。然而，由於K-Means的相關參數和閾值都是使用初始設定而沒有進一步動態調整，使其推薦結果的正確率改進幅度不大(曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯，2005)。

(四) 使用主題地圖之視覺化滿意度

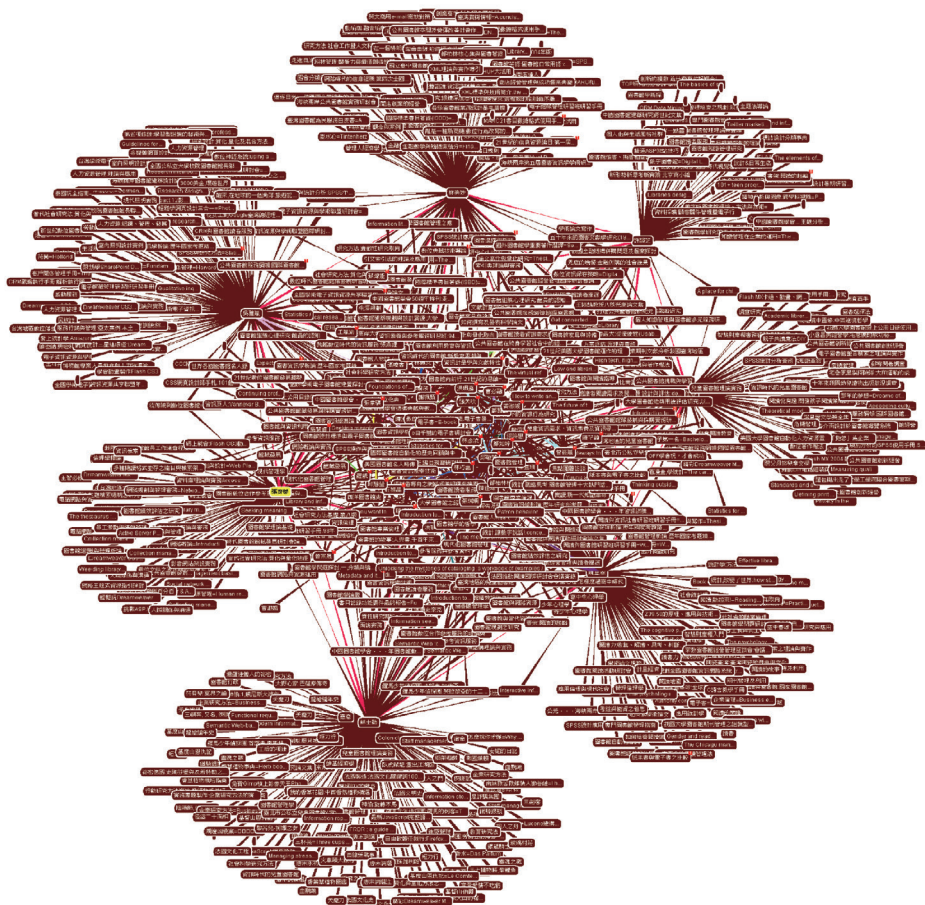
由於傳統圖書推薦系統無法告知使用者推薦的原因(Verbert et al., 2013)，藉由主題地圖的優越顯示功能，除可讓使用者一目了然圖書推薦結果的隱含結構性，更能有效利用推薦結果。因此，本研究提出單層式和雙層式主題地圖呈現。為了解這兩種視覺化圖書推薦結果呈現的效能，使用如附錄一所示的讀者滿意度問卷。此問卷為李克特式五分量表。

1. 單層式主題地圖呈現

單層式主題地圖呈現採用的是非常直覺性的「人-借-書」的概念來建立關聯。此呈現方式希望讓讀者能一眼看出與自己相同群組的其他讀者還借過什麼書，除可以了解推薦系統之推薦原因，還可自行推導系統無法推薦的圖書。然而，由於圖書推薦系統所處理圖書的數量非常龐大且書名通常很長，若全部以主題地圖呈現的話，會如示意圖11所示，互相重疊顯示在一起，使讀者難以直覺辨識反而無法利用。

2. 雙層式主題地圖呈現

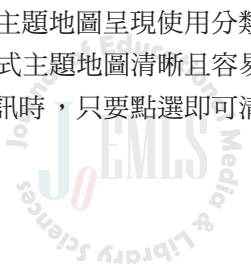
加入分類的概念，改由「人-借-分類」和「分類-包含-書」的雙層連結架構，以減少與讀者這個主題第一層連結的主題數量和縮減整體的主題顯示空間，同時也提升了「人-借-書」概念的層次性。另一方面，為了改進原本幾乎是同一

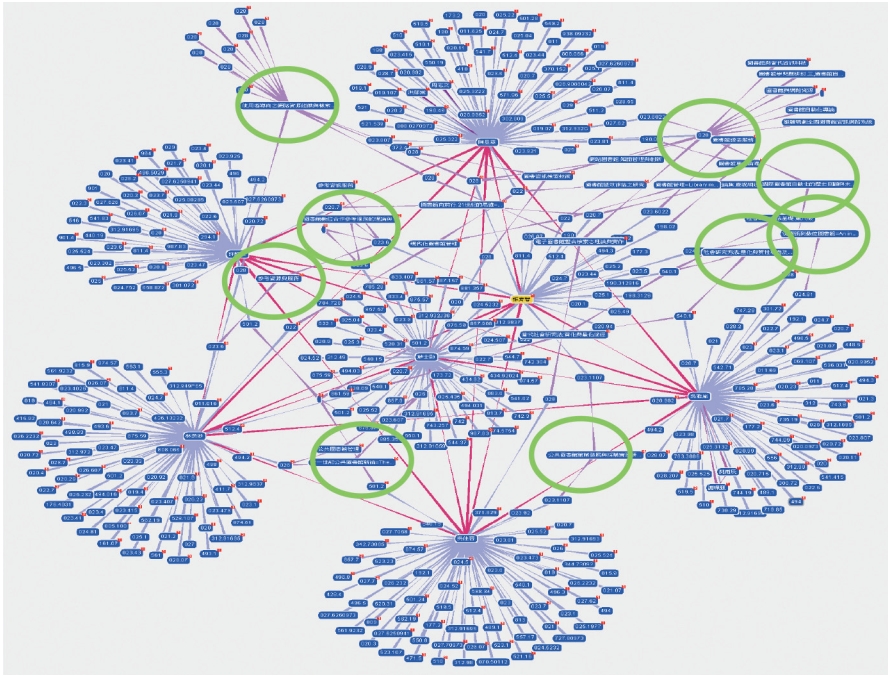


(彩圖請見電子檔)

圖11 單層式主題地圖呈現

顏色涵蓋所有主題和關聯的情況，提高主題與關聯之間的辨識度，增加了許多顏色上的設定。原則是讓不同類型的關聯能以不同顏色呈現，但同類主題都採用同一顏色，以免整體顏色過多反造成視覺辨識的困難度。雙層式主題地圖呈現示意圖如圖12所示，藍色圓角方框是主題，紅色線條是群組關聯，淺藍色線條是借閱關聯，紫色線條代表包含關聯。由於雙層式主題地圖呈現使用分類號取代書名，由於分類號的號碼較短，顯示效果比單層式主題地圖清晰且容易了解。一旦讀者需要了解此分類號的書名或其他相關資訊時，只要點選即可清楚顯示。





(彩圖請見電子檔)

圖12 雙層式主題地圖呈現

3. 實驗結果

由前述四位圖資所研究生挑選無使用圖形化呈現經驗的三位同學，經過簡短測試目的和問卷內容說明後，使用附錄一的問卷進行主題地圖呈現滿意度調查(如表3所示)。除去第7題和第8題，以五分尺度計算，非常同意為5分，非常不同意為1分。實驗一(單層式)和實驗二(雙層式)的平均分數分別為2.9和3.8外，問題1、2、9、10的結果皆表示雙層式主題地圖呈現確實可讓使用者獲

表3 主題地圖呈現滿意度問卷結果

題號	單層式主題地圖			雙層式主題地圖		
	A生	B生	C生	A生	B生	C生
1	4	4	2	4	4	3
2	4	3	3	4	5	3
3	2	2	1	4	5	4
4	1	3	1	2	3	3
5	4	4	4	5	5	4
6	2	4	3	4	4	3
7	5-6本	3-4本	3-4本	9本以上	9本以上	7-8本
8	2組	2組	1組	2組	2組	2組
9	3	5	2	4	5	4
10	3	4	1	4	4	4
平均	2.9	3.7	2.1	3.8	4.1	3.5
總平均		2.9			3.8	

得更大滿意度。依據問題3的結果，由於雙層式結果明顯優於單層式，顯示使用者不希望第一次即呈現所有的資訊，有需要時再自行點選呈現即可。這和人機介面研究的使用者控制 (User control) 概念完全一致。由於書名通常具有一定長度，單層式主題地圖呈現時佔有很大顯示空間。當讀者借閱書籍多時，以讀者為中心的主題圈會顯得非常密集、雜亂。個人借閱已是如此，更遑論所有屬性同時顯現的情形。相反的，雙層式主題地圖呈現經過層次和顏色的區別後，可明顯看出讀者與讀者間書籍借閱的關聯性。在第二層次展開的情況下，同一群組內讀者借閱過相同書籍也會被顯示 (如綠色圓圈標示)。若需要更進一步推薦讀者其他同群組讀者借過的書籍，則可向下展開第三層次，展開個別/群組/所有讀者所有借過的書籍。因此，雙層式主題地圖呈現比單層式主題地圖呈現更為友善，易於被讀者接受。

接著，由問題4的結果發現，上述兩種主題地圖呈現都有大量資料所造成反應速度過慢的問題，分析其原因為隨著資料量的增加，主題地圖轉換模組和主題地圖後處理模組的處理時間也會隨著加大，造成使用者等待時間的心裡負擔。如何縮短等待時間，例如只顯示更精簡的資訊等，將是實用化時必須解決的課題。再者，由問題5、6之可否縮短搜尋館藏資源時間和是否增加搜尋資源便利性的結果，雖然雙層式主題地圖呈現和單層式主題地圖呈現在搜尋館藏資源時間上有明顯的縮短；然而在是否增加搜尋資源便利性上卻無明顯的改進。由圖12所示，雙層式主題地圖呈現仍有過多的重複，例如左上角都是028，右上角也有028對應對到多本書籍，其原因可能是仍有過多的分類號，造成無法增加搜尋資源的便利性。另一方面，由問題7的結果，可發現單層式主題地圖呈現太多資訊，造成使用者不易閱讀而希望呈現較少資訊 (3-4本等)，然而由於雙層式主題地圖呈現較少資訊，故使用者希望多呈現資訊 (9本以上)。由問題8測試結果，也顯示雙層式主題地圖呈現仍不夠精簡，造成使用者如同使用單層式主題地圖呈現同樣不希望呈現太多群組 (兩組即可)。

五、結 論

為導入 Web2.0 概念至圖書館服務，本研究提出時間衰減因素、導入矩陣分群和主題地圖視覺化圖書推薦系統。其中，為解決高維度和稀疏二值矩陣所造成的分群問題，導入了矩陣分群演算法、動態閾值和最大密實子矩陣的評價函數後，使用貪婪式搜尋法使得可在合理時間內得到接近最佳解的分群結果。另一方面，為讓讀者可一目了然了解與其他讀者或圖書之間的關係，可利用關係推導進而發現深層的隱藏知識。進一步導入以圖形呈現讀者-讀者、讀者-圖書，或圖書-圖書等關聯的主題地圖，並探討適合讀者的主題地圖呈現方式。由實驗結果顯示，本研究提出之視覺化圖書推薦系統，比傳統分群方法可

得到更適切的分群結果外，雙層式主題地圖呈現比單層式主題地圖更清楚明了的讀者與圖書之間的關係。為提高視覺化圖書推薦系統的效率，今後仍有一些可以繼續研究的課題，分別敘述如下：

(一)對於冷啟動和個人隱私等本研究尚未解決的課題藉著導入社群網路 (social network) 方可解決 (Huang, Contractor, & Yao, 2008)。再者，可導入關係推導演算法或資料探勘技術，由視覺化網路社群推薦結果中挖掘深層的知識架構和表徵 (Guha, Kumar, Raghavan, & Tomkins, 2004; Jung, 2005)。

(二)本研究選擇的讀者都是圖書資訊學研究所的研究生。這些學生由於修課與研究領域有很大的重疊性或同值性，且實驗資料只使用59人，實驗的有效性無法證實是明顯的 (Significant)。為進一步證實本研究的有效性，未來應加大實驗資料的數量，例如隨機由圖書館讀者資料中隨機挑選非同值性的讀者資料加入實驗，進一步交叉比對與討論本研究的有效性。

(三)使用圖書館的讀者借還書記錄，評估本視覺化圖書推薦系統的求全率 (recall rate) 與求準率 (precision rate)，以進一步提高系統的效率。

(四)問卷的結果顯示主題地圖的視覺化圖書推薦系統確實可提供使用者互動功能，也可提供讀者理解系統推薦圖書的原因。進一步可針對系統的使用性 (usability) 和有用性 (usefulness)，探討各式各樣主題地圖的呈現方式，提供讀者容易操作的人機介面。另外，本研究之主題地圖呈現使用「中國圖書分類號」，對於非圖書資訊專業人員而言，需要額外說明才可看懂。主題地圖的有效呈現 (例如，不使用分類號，只顯示大類的說明等) 也是一個重要課題。

(五)分群可導入分散式或多執行緒等運算，有效解決大量資料運算速度的課題。然而，視覺化呈現隨著資料量的增加，大量XML資料轉換產生運算緩慢的問題。因此，如何提高圖形化呈現效率將是一個重要的研究課題，例如，合併類似項目或只顯示最相關者。

(六)參考統計學的趨勢分析技術，導入其他時間衰退因素的影響要素，進一步提高推薦系統的效能。

誌 謝

感謝國科會計畫 NSC 99-2221-E-005-097 對於本研究的經費補助和支持，並承蒙國立中興大學圖書館提供相關資料並給予建議與協助，特此誌謝。

參考文獻

- 陳垂呈、陳幸暉 (2011)。建置圖書館書籍推薦系統：資料探勘之應用。《工程科技與教育學刊》，8(3)，469-478。
- 卜小蝶 (1998年10月)。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》，2，63-73。

- 郭逸凡 (2003)。以矩陣分群技術分析顧客行為模式 (未出版之碩士論文)。國立成功大學資訊管理研究所，台南市。
- 賴永祥 (編訂) (2001)。中國圖書分類法 (增訂八版)。台北市：文華。
- 曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯 (2005)。資料探勘。台北市：旗標。
- Aciar, S., Zhang, D., Simoff, S., & Debenham, J. (2007). Informed recommender: Basing recommendations on consumer product reviews. *IEEE Intelligent Systems*, 22(3), 39-47. doi:10.1109/MIS.2007.55
- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749. doi:10.1109/TKDE.2005.99
- Angiulli, F., Cesario, E., & Pizzuti, C. (2006). A greedy search approach to co-clustering sparse binary matrices. In S. Ceballos (Ed.), *Tools with artificial intelligence, 2006. ICTAI '06. 18th IEEE international conference on* (pp. 363-370). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/ICTAI.2006.10
- Chen, Z., Jiang, Y., & Zhao, Y. (2010). A collaborative filtering recommendation algorithm based on user interest change and trust evaluation. *International Journal of Digital Content Technology and its Applications*, 4(9), 106-113. doi:10.4156/jdcta.vol4.issue9.13
- CSVParser. (2011). Retrieved from <http://ostermiller.org/utills/doc/com/Ostermiller/utill/CSVParser.htm>
- Gong, S. J., & Cheng, G. H. (2008). Mining user interest change for improving collaborative filtering. In R. Bilof (Ed.), *Proceeding IITA '08 proceedings of the 2008 second international symposium on intelligent information technology application* (Vol. 3, pp. 24-27). Washington, DC: IEEE Computer Society. doi:10.1109/IITA.2008.385
- Guha, R., Kumar, R., Raghavan, P., & Tomkins, A. (2004). Propagation of trust and distrust. In S. Feldman (Ed.), *Proceeding WWW '04 proceedings of the 13th international conference on World Wide Web* (pp. 403-412). New York, NY: ACM. doi:10.1145/988672.988727
- Hwang, S.-Y., & Lim, E.-P. (2002). A data mining approach to library new book recommendations. In E.-P. Lim, S. Foo, C. Khoo, H. Chen, E. Fox, S. Urs, & T. Costantino (Eds.), *Lecture Notes in Computer Science: Vol. 2555. Digital libraries: People, knowledge, and technology: 5th international conference on Asian digital libraries, ICADL 2002, Singapore, December 11-14, 2002: proceedings* (pp. 229-240). New York, NY: Springer. doi:10.1007/3-540-36227-4_23
- Huang, Y., Contractor, N., & Yao, Y. (2008). CI-KNOW: Recommendation based on social networks. In S. A. Chun, M. Janssen, & J. R. Gil-Garcia (Eds.), *Proceeding dg.o '08 proceedings of the 2008 international conference on digital government research* (pp. 27-33). Montreal, Canada: Digital Government Society of North America.
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Survey*, 31(3), 264-323. doi:10.1145/331499.331504
- Jung, J. J. (2005). Visualizing recommendation flow on social network. *Journal of Universal Computer Science*, 11(11), 1780-1791. doi:10.3217/jucs-011-11-1780
- Kuo, J. J., & Zhang, Y. J. (2012). A library recommender system using interest change over time and matrix clustering. In H.-H. Chen & G. Chowdhury (Eds.), *Lecture Notes in Computer*

- Science: Vol. 7634. The outreach of digital libraries: A globalized resource network: 14th international conference on Asia-Pacific digital libraries, ICADL 2012, Taipei, Taiwan, November 12-15, 2012, proceedings* (pp. 259-268). New York, NY: Springer Berlin Heidelberg. doi:10.1007/978-3-642-34752-8_32
- Kumar, B. R. S., Ratnam, B. J., & Babu, M. S. P. (2010). Improvement of personalized recommendation algorithm based on hybrid collaborative filtering. *International Journal of Computer Science & Communication*, 1(2), 429-432.
- Kuroiwa, T., & Bhalla, S. (2007). Dynamic personalization for book recommendation system using web services and virtual library enhancements. In T. Miyazaki, I. Paik, & D. Wei (Eds.), *Computer and information technology, 2007. CIT 2007. 7th IEEE international conference on* (pp. 212-217). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/CIT.2007.72
- Maness, J. M. (2006). Library 2.0 theory: Web 2.0 and its implications for libraries. *Webology*, 3(2), Article 25. Retrieved from <http://www.webology.org/2006/v3n2/a25.html>
- McLaughlin, M. R., & Herlocker, J. L. (2004). A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In K. Järvelin (Ed.), *Special Interest Group on Information Retrieval: Vol. '04. Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 329-336). New York, NY: ACM. doi: 10.1145/1008992.1009050
- TopicMaps.Org. (2005). *XML Topic Maps (XTM) v1.0*. Retrieved from <http://www.topicmaps.org/xtm/1.0/>
- Mooney, R. J., & Roy, L. (2000). Content-based book recommending using learning for text categorization. In *ACM 2000 digital libraries: Proceedings of the fifth ACM conference on digital libraries, June 2-7, 2000, San Antonio, Texas, USA* (pp. 195-204). New York, NY: ACM. doi:10.1145/336597.336662
- Omniator. (2009). Retrieved from <http://www.ontopia.net/download/index.html>
- Oyanagi, S., Kubota, K., & Nakase, A. (2001). Matrix clustering: A new data mining algorithm for CRM. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, 100(351), 25-32.
- Oyanagi, S., Kubota, K., & Nakase, A. (2001, August). Application of matrix clustering to web log analysis and access prediction. In R. Bayardo (Chair) *WEBKDD 2001-mining web log data across all customers touch points*. WEBKDD, San Francisco, CA.
- Peis, E., del Castillo, J. M. M., & Delgado-Lopez, J. A. (2008). Semantic Recommender Systems. Analysis of the state of the topic. *Hipertext.net*, 6. Retrieved from <http://www.upf.edu/hipertextnet/en/numero-6/recomendacion.html>
- Pepper, S. (2005). *The TAO of topic maps: Finding the way in the age of Infoglut*. Retrieved from <http://www.ontopia.net/topicmaps/materials/tao.html>
- Resnick, P., Lacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. In R. Furuta & C. Neuwirth (Eds.), *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp.175-186). New York, NY: ACM. doi:10.1145/192844.192905

- Ricci, F., Rokach, I., Shapira, B., & Kantor, P. B. (Eds.). (2011). *Recommender system handbook*. New York, NY: Springer.
- Russell, S., & Norvig, P. (2003). *Artificial intelligence: A modern approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Verbert, K., Parra, D., Brusilovsky, P., & Duval, E. (2013). Visualizing recommendations to support exploration, transparency and controllability. In J. Kim, J. Nichols, & P. A. Szekely (Eds.), *IUI '13 companion: Proceedings of the companion publication of the 2013 international conference on intelligent user interfaces companion* (pp. 351-362). New York, NY: ACM. doi:10.1145/2449396.2449442
- Zhang, Q. Q., & Ye, N. (2012). Collaborative Filtering Algorithm Adapting to Changes Over Dynamic Time. In L. Xu (Ed.), *International Proceedings of Computer Science and Information Technology: Vol. 30. Network and computer science: Selected, peer reviewed papers from the 2012 international conference on network and computer science (ICNCS 2012), March 10-11, 2012, Hong Kong* (pp. 27-32). Singapore: IACSIT Press.

附錄一 以主題地圖呈現圖書推薦之滿意度調查問卷範例

1. 您認為主題地圖圖書推薦系統的介面操作容易？
非常同意 同意 普通 不同意 非常不同意
2. 您認為主題地圖圖書推薦系統的功能設計簡易？
非常同意 同意 普通 不同意 非常不同意
3. 您認為主題地圖圖書推薦系統的畫面呈現清楚？
非常同意 同意 普通 不同意 非常不同意
4. 您認為主題地圖圖書推薦系統的反應速度過慢？
非常同意 同意 普通 不同意 非常不同意
5. 您認為主題地圖圖書推薦能縮短您搜尋相關館藏資源的時間？
非常同意 同意 普通 不同意 非常不同意
6. 您認為主題地圖圖書推薦能增加您搜尋相關館藏資源的便利性？
非常同意 同意 普通 不同意 非常不同意
7. 您認為主題地圖圖書推薦畫面上應呈現多少本圖書較適當？
0-2本 3-4本 5-6本 7-8本 9本以上
8. 您認為主題地圖圖書推薦系統的顯示應為多少群組較適合？
1組 2組 3組 4組 5組以上
9. 您樂意繼續使用主題地圖圖書推薦系統？
非常同意 同意 普通 不同意 非常不同意
10. 整體而言，您對主題地圖圖書推薦感到滿意？
非常同意 同意 普通 不同意 非常不同意



Visualized Book Recommender System Using Matrix Clustering

June-Jei Kuo*

Assistant Professor
E-mail: jjkuo@dragon.nchu.edu.tw

Jui-Shan Chang

Graduate Student
E-mail: changrs.tw@gmail.com

Yu-Jung Zhang

Graduate Student
Graduate Institute of Library and Information Science
National Chung Hsing University
Taipei, Taiwan
E-mail: tatina8088vs@yahoo.com.tw

Abstract

Traditional library recommender system can not only employ users' borrowing records to recommend books with similar subjects which they have read, but also use borrowing records of users who are in the same social network to recommend books they never borrow but may be interested in. However, as users' reading interests changes from time to time, treating their borrowing records at different time periods equally seems to lead the recommendation results not to meet the users' current needs. Moreover, as the borrowing records are highly dimensional and sparse, the traditional clustering methods cannot tackle clustering issue effectively. Besides, in order to allow users to examine recommendation results in multiple aspects and offer a clear picture of items ranked by users' perceived reading interests, interactive information visualization need to be implemented. Therefore, this paper exploits time decay weight, matrix clustering using dynamic threshold and topic maps to propose a novel visualized book recommender system. According to the experimental results of users' satisfaction questionnaire, the proposed recommender system can be useful to represent the recommendation results and helpful for users to find their interested books. Furthermore, two-layered topic map is easier to understand than one-layered topic map, and it can effectively satisfy the users' needs.

Keywords: Time decay; Matrix clustering; Topic map; Book recommender

* Principal author for all correspondence.

SUMMARY

1. Introduction

With the challenge of information explosion and development of Web 2.0 technology, traditional libraries have transformed into digital libraries gradually. The abundance of documents and information in digital libraries has created much inconvenience to users seeking information online. Consequently, digital libraries are confronted with a big challenge: effective management of large information collections to give users easy and quick access to information needed. Along with personalized service development and research, recommender systems have proven to be valuable in tackling the above issues advising users on relevant products and information by predicting a user’s interest, based on various types of information. Currently, collaborative filtering has been one of the most successful technologies for personalized recommendation, in which the typical algorithm is the clustering algorithm using similarity based on user ratings in order to find neighbors who have similar interests with the active user. However, most of the traditional collaborative filtering algorithms only pay attention to the similarity score, without considering the user interest shifting with the change of time. Thus leads to the recommendation result departure from the true user’s current information needs. In addition, for library recommendation application, high dimensional and sparse data are typical. Besides, interactive information visualization can allow users to more easily see multiple aspects of recommended results and offer a clear of items ranked by perceived interests. To solve the above issues of library recommender systems simultaneously, the time decay weight for interest change over time, matrix clustering and topic maps are introduced to propose a novel library recommender system.

2. Matrix Clustering

To deal with the issue of personalization, how to obtain customers’ interests effectively has become an important research topic. One of the approaches for personalization is matrix clustering, which is similar to the collaborative filtering. Figure 1 shows the exchange example of rows and columns and the extracted dense sub-matrix from the given base matrix.

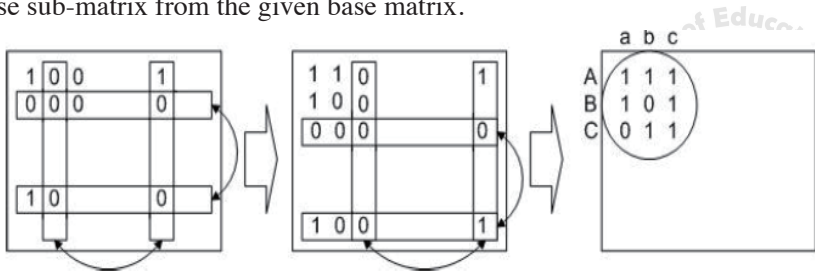


Figure 1 Exchange Example of Rows and Columns

A naive algorithm to exchange rows and columns iteratively requires many computations when the matrix is large. Ping-Pong algorithm is a new fast algorithm, which is intended to reduce the execution time by utilizing the sparseness of a matrix. Furthermore, the Ping-Pong algorithm finds the most maximum dense matrix, which satisfies both the support and confidence. However, we want to extract all the dense matrices iteratively. Thus, the Ping-Pong algorithm is modified below.

Matrix Clustering using Dynamic Parameters:

- (1) Input binary sparse matrix A and set the related parameters such as density.
- (2) Use random number to decide the start row and then activate the row in matrix A.
- (3) Execute the ping-pong algorithm.
- (4) Delete the rows and columns of dense matrix from matrix A.
- (5) If there is an extracted dense matrix, output the dense matrix and go to step 2
- (6) If matrix A is not NULL, decrease the density parameter and activated value stepwise.
- (7) If both the density and activated value are bigger than predefined respective critical values, go to step 2.
- (8) End

3. Visualized Book Recommender System

The architecture of proposed visualized library recommender system is shown as Figure 2. There are six modules and three types of records collected by digital libraries. Besides, the borrow/return statistics is created by data conversion module.

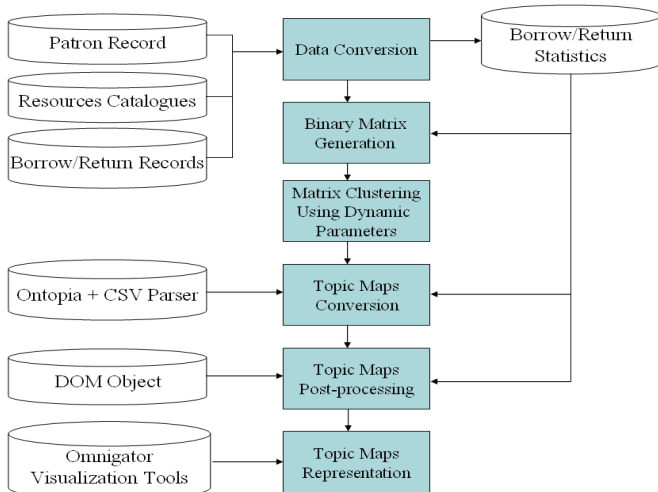


Figure 2 Architecture of the Proposed Visualized Library Recommender System

(1) Data Conversion Module

The module deletes the unnecessary data and saves the remainder. All the data belongs to a user will first be gathered and then the relevant data of each user are merged into borrow/return statistics. Moreover, the start date, return date,

loan duration, and borrowing times of each book are computed and saved as well.

(2) Binary Matrix Generation Module

This module converts the data from data conversion module into binary matrix using some criteria. Moreover, the time weight of user access time is also introduced to improve the importance of recent borrowed books in the course of producing the book recommendation.

(3) Matrix Clustering Using Dynamic Parameters Module

This module employs the matrix clustering to detect maximal and dense sub-matrices. As the matrix clustering is an NP problem, an admissible heuristic function is needed. In order to obtain a heuristic function, the matrix clustering problem is eased with fewer restrictions.

(4) Topic Maps Generation Module

To convert the data into a topic map, first, users' interesting group data are automatic clustered by the dynamic threshold matrix clustering module. Then, we collect the users' profiles, circulation records and library bibliographic records and convert all data into CSV files. The conversion module is developed in a JAVA environment using the Topic Maps Engine and Ontopia API to build the topic map. In order to read the file content in CSV format, we make use of the CSVParser package to parse the value in the column field. Finally, a topic map with XTM format is generated.

(5) Topic Maps Post-processing Module

There is a serious representing problem with the topic map generated by the conversion module: the topics of the same concept are not identified as the same object in the topic map. To cope with this problem, the DOM(Document Object Model) package is used. An ID list was built to process the duplicate objects. Topics with wrong reference will be searched and then deleted.

(6) Topic Maps Representation Module

Omnigator is a topic map navigator built using components of the Ontopia for loading, viewing, visualizing, debugging and merging Topic Maps. For this reason, we not only build the topic maps with the Ontopia Topic Maps Engine and API, but also visualize the topic maps with Omnigator to make the presentation more intuitive and effective for browsing and evaluating.

4. Experiment Results

The related borrow/return records of our graduate students from year 2005 to 2010 were employed to be our test corpus. There were a total of 71 students. After deleting those students who seldom or never borrowed books, there were 59 students. Furthermore, the related patron records and resources catalogues were extracted as well.

(1) Results of the Satisfaction of Recommender Systems

To compare the proposed library recommender systems, K-Means clustering (baseline1) and hierarchical clustering using complete link algorithm (baseline2) are employed to implement two baseline recommender systems. The results are shown in Table 1, the scores are precisions. The higher the precisions are, the better the recommendation performance is. Moreover, time decay A and B means “not using time weight” and “using time weight”. On the one hand, Rigid metric used only the marks in the “interested” column of evaluation sheets. On the other hand, the relaxed metrics used all the marks in the “interested” or “don’t know” columns of evaluation sheets.

Table 1 Average Precision of Experimental Results

Time decay	Methods		Baseline1	Baselin2	Proposed Method
	Metrics				
A	Rigid		0.2918	0.2755	0.3075
	Relaxed		0.3510	0.3320	0.6294
B	Rigid		0.3350	0.3725	0.5161
	Relaxed		0.3817	0.4089	0.7124

(2) Results of Visualized Satisfaction

Two kinds of Topic Map presentations are proposed and named as one-layered topic map and two layered topic, respectively. For each topic node, one-layered topic map shows all the information. On the other hand, for each topic node the two-layered topic map hides most of information and only shows the necessary information. Whenever necessary, the readers can expand the information easily. To understand the usability and visibility of the proposed Topic Map presentations, five-point Likert scale questionnaire and three testers are employs to evaluate the visualized satisfaction. The average scores for both topic map presentations are 2.9 (one-layered) and 3.8 (two-layered), respectively. Apparently, two-layered topic map is more easily understood than one-layered topic map, and can effectively satisfy the users’ information needs.

5. Conclusions

During the time of knowledge explosion, personalized knowledge service is an inevitable trend and it exhibits a high-level service pattern to satisfy users’ diversified information needs. The study integrates users’ interest change over time into user interest modeling. Additionally, to tackle the data sparsity the matrix clustering is also introduced and modified to discover neighbors who have same interests. Topic map is also employed to propose a novel visualized recommender system. According to the experiment promising results, the time decay weight can effectively reflect user’s interest trend and the matrix clustering

can tackle the sparsity issue as well. Moreover, for better visibility and usability, the two-layered topic map presentation is superior than one-layered topic map presentation. To further improve the quality of the proposed library recommender system there are future works described below.

- (1) To cope with the scalability problem, effective heuristic functions should be further studied.
- (2) How to provide a user-friendly interface such as semantic web or topic map is indispensable.
- (3) To enhance the performance of the recommender system, social network and trust propagation should be researched and introduced.

ROMANIZED & TRANSLATED REFERENCE FOR ORIGINAL TEXT

陳垂呈、陳幸暉 [Chen, Chui-Cheng, & Chen, Hsing-Hui] (2011)。建置圖書館書籍推薦系統：資料探勘之應用 [Jianzhi tushuguan shuji tuijian xitong: Ziliao tankan zhi yingyong]。工程科技與教育學刊 [Journal of Engineering Technology and Education], 8(3), 469-478。

卜小蝶 [Pu, Hsiao-Tieh] (1998年10月) [(1998, October)]。淺析個人化服務技術的發展趨勢對圖書館的影響 [Gianxi gerenhua fuwu jishu de fazhan qushi dui tushuguan de yingxiang]。國立成功大學圖書館館刊 [National Cheng Kung University Library Journal], 2, 63-73。

郭逸凡 [Kuo, Evan] (2003)。以矩陣分群技術分析顧客行為模式 (未出版之碩士論文) [Yi juzhen fenqun jishu fenxi guke xingwei moshi (Unpublished master's thesis)]。國立成功大學資訊管理研究所，台南市 [Graduate Institute of Information Management, National Cheng Kung University, Tainan]。

賴永祥 (編訂) [Lai, Yung-Hsiang Johy (Ed.)] (2001)。中國圖書分類法 (增訂八版) [New classification scheme for Chinese libraries (Rev. 8th ed.)]。台北市：文華 [Taipei: Mandarin Library & Information Services]。

曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯 [Tseng, Sian-Syong, Tsai, Hsiu-Man, Su, Tung-Hsing, Tseng, Chiu-Jung, & Wang, Ching-Yao] (2005)。資料探勘 [Data Mining]。台北市：旗標 [Taipei: Flag Publishing]。

Acıar, S., Zhang, D., Simoff, S., & Debenham, J. (2007). Informed recommender: Basing recommendations on consumer product reviews. *IEEE Intelligent Systems*, 22(3), 39-47. doi:10.1109/MIS.2007.55

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749. doi:10.1109/TKDE.2005.99

Angiulli, F., Cesario, E., & Pizzuti, C. (2006). A greedy search approach to co-clustering sparse binary matrices. In S. Ceballos (Ed.), *Tools with artificial intelligence, 2006. ICTAI '06. 18th IEEE international conference on* (pp. 363-370). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/ICTAI.2006.10

- Chen, Z., Jiang, Y., & Zhao, Y. (2010). A collaborative filtering recommendation algorithm based on user interest change and trust evaluation. *International Journal of Digital Content Technology and its Applications*, 4(9), 106-113. doi:10.4156/jdcta.vol4.issue9.13
- CSVParser. (2011). Retrieved from <http://ostermiller.org/utills/doc/com/Ostermiller/util/CSVParser.htm>
- Gong, S. J., & Cheng, G. H. (2008). Mining user interest change for improving collaborative filtering. In R. Bilof (Ed.), *Proceeding IITA '08 proceedings of the 2008 second international symposium on intelligent information technology application* (Vol. 3, pp. 24-27). Washington, DC: IEEE Computer Society. doi:10.1109/IITA.2008.385
- Guha, R., Kumar, R., Raghavan, P., & Tomkins, A. (2004). Propagation of trust and distrust. In S. Feldman (Ed.), *Proceeding WWW '04 proceedings of the 13th international conference on World Wide Web* (pp. 403-412). New York, NY: ACM. doi:10.1145/988672.988727
- Hwang, S.-Y., & Lim, E.-P. (2002). A data mining approach to library new book recommendations. In E.-P. Lim, S. Foo, C. Khoo, H. Chen, E. Fox, S. Urs, & T. Costantino (Eds.), *Lecture Notes in Computer Science: Vol. 2555. Digital libraries: People, knowledge, and technology: 5th international conference on Asian digital libraries, ICADL 2002, Singapore, December 11-14, 2002: proceedings* (pp. 229-240). New York, NY: Springer. doi:10.1007/3-540-36227-4_23
- Huang, Y., Contractor, N., & Yao, Y. (2008). CI-KNOW: Recommendation based on social networks. In S. A. Chun, M. Janssen, & J. R. Gil-Garcia (Eds.), *Proceeding dg.o '08 proceedings of the 2008 international conference on digital government research* (pp. 27-33). Montreal, Canada: Digital Government Society of North America.
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Survey*, 31(3), 264-323. doi:10.1145/331499.331504
- Jung, J. J. (2005). Visualizing recommendation flow on social network. *Journal of Universal Computer Science*, 11(11), 1780-1791. doi:10.3217/jucs-011-11-1780
- Kuo, J. J., & Zhang, Y. J. (2012). A library recommender system using interest change over time and matrix clustering. In H.-H. Chen & G. Chowdhury (Eds.), *Lecture Notes in Computer Science: Vol. 7634. The outreach of digital libraries: A globalized resource network: 14th international conference on Asia-Pacific digital libraries, ICADL 2012, Taipei, Taiwan, November 12-15, 2012, proceedings* (pp. 259-268). New York, NY: Springer Berlin Heidelberg. doi:10.1007/978-3-642-34752-8_32
- Kumar, B. R. S., Ratnam, B. J., & Babu, M. S. P. (2010). Improvement of personalized recommendation algorithm based on hybrid collaborative filtering. *International Journal of Computer Science & Communication*, 1(2), 429-432.
- Kuroiwa, T., & Bhalla, S. (2007). Dynamic personalization for book recommendation system using web services and virtual library enhancements. In T. Miyazaki, I. Paik, & D. Wei (Eds.), *Computer and information technology, 2007. CIT 2007. 7th IEEE international conference on* (pp. 212-217). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/CIT.2007.72
- Maness, J. M. (2006). Library 2.0 theory: Web 2.0 and its implications for libraries. *Webology*, 3(2), Article 25. Retrieved from <http://www.webology.org/2006/v3n2/a25.html>

- McLaughlin, M. R., & Herlocker, J. L. (2004). A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In K. Järvelin (Ed.), *Special Interest Group on Information Retrieval: Vol. '04. Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 329-336). New York, NY: ACM. doi: 10.1145/1008992.1009050
- TopicMaps.Org. (2005). *XML Topic Maps (XTM) v1.0*. Retrieved from <http://www.topicmaps.org/xtm/1.0/>
- Mooney, R. J., & Roy, L. (2000). Content-based book recommending using learning for text categorization. In *ACM 2000 digital libraries: Proceedings of the fifth ACM conference on digital libraries, June 2-7, 2000, San Antonio, Texas, USA* (pp. 195-204). New York, NY: ACM. doi:10.1145/336597.336662
- Omnigator. (2009). Retrieved from <http://www.ontopia.net/download/index.html>
- Oyanagi, S., Kubota, K., & Nakase, A. (2001). Matrix clustering: A new data mining algorithm for CRM. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, 100(351), 25-32.
- Oyanagi, S., Kubota, K., & Nakase, A. (2001, August). Application of matrix clustering to web log analysis and access prediction. In R. Bayardo (Chair) *WEBKDD 2001-mining web log data across all customers touch points*. WEBKDD, San Francisco, CA.
- Peis, E., del Castillo, J. M. M., & Delgado-Lopez, J. A. (2008). Semantic Recommender Systems. Analysis of the state of the topic. *Hipertext.net*, 6. Retrieved from <http://www.upf.edu/hipertextnet/en/numero-6/recomendacion.html>
- Pepper, S. (2005). *The TAO of topic maps: Finding the way in the age of Infoglut*. Retrieved from <http://www.ontopia.net/topicmaps/materials/tao.html>
- Resnick, P., Lacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. In R. Furuta & C. Neuwirth (Eds.), *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp.175-186). New York, NY: ACM. doi:10.1145/192844.192905
- Ricci, F., Rokach, I., Shapira, B., & Kantor, P. B. (Eds.). (2011). *Recommender system handbook*. New York, NY: Springer.
- Russell, S., & Norvig, P. (2003). *Artificial intelligence: A modern approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Verbert, K., Parra, D., Brusilovsky, P., & Duval, E. (2013). Visualizing recommendations to support exploration, transparency and controllability. In J. Kim, J. Nichols, & P. A. Szekely (Eds.), *IUI '13 companion: Proceedings of the companion publication of the 2013 international conference on intelligent user interfaces companion* (pp. 351-362). New York, NY: ACM. doi:10.1145/2449396.2449442
- Zhang, Q. Q., & Ye, N. (2012). Collaborative Filtering Algorithm Adapting to Changes Over Dynamic Time. In L. Xu (Ed.), *International Proceedings of Computer Science and Information Technology: Vol. 30. Network and computer science: Selected, peer reviewed papers from the 2012 international conference on network and computer science (ICNCS 2012), March 10-11, 2012, Hong Kong* (pp. 27-32). Singapore: IACSIT Press.

