

應用模糊語意法找尋協同式篩選系統最近鄰居之研究

Searching the Nearest Neighbor in Collaborative Filtering Systems using the Fuzzy Linguistic Analysis

李來錫¹ 蔡耀德²

(Received: Aug. 14, 2006 ; First Revision: Nov. 1, 2006 ; Accepted: Nov. 20, 2006)

摘要

協同式篩選系統是目前常用的資訊篩選方法之一，此系統會根據有相同興趣的成員的資訊作為推薦參考項目。此系統必須辨識相近的使用者，即找出最近鄰居，因此如何有效進行最近鄰居的挑選，是協同式篩選系統的首要課題。本研究目的在以模糊語意法進行協同式篩選系統之最近鄰居的尋找工作。在尋求最近鄰居的做法上，首先在篩選系統設置資訊涉入問卷，並要求每位新進使用人進行填寫，以收集所有使用者對於系統提供資訊的涉入狀況。一旦有目標使用人欲進行最似鄰居的挑選時，系統首先以模糊語意法找出目標使用人與鄰居間的感受，再進行兩者間的離異度計算。最後，本研究以實例進行最近鄰居的挑選，並順利找出最近鄰居。研究結果顯示，模糊語意法可以提供較細緻最近鄰居挑選程序，並獲得良好的成效。

關鍵詞：模糊語意、協同式篩選系統、最近鄰居

Abstract

Finding the nearest neighbor becomes the main topic for designing the collaborative filtering system. This study is aim to search the nearest neighbor in collaborative filtering systems using the fuzzy linguistic analysis. In the beginning of the filtering system a questionnaire of information involvement is set to collect data for new users. Then the fuzzy linguistic analysis is used to analysis the difference of apprehension among users in the questionnaire. Additionally, a difference index between two users is then calculated as the comparison base for searching the nearest neighbor. An example is used to illustrate the procedure of the method and show the possibility for searching the nearest neighbor.

Keywords : fuzzy linguistic analysis, collaborative filtering system, nearest neighbor

¹國立屏東商業技術學院資訊管理系助理教授

²國立屏東商業技術學院資訊管理系研究生

1. 緒論

隨著網際網路(internet)的快速發展，使用者在虛擬的網路空間(cyberspace)中彼此共享資訊或透過相關工具取得所需資訊。然而，網路科技資訊的供給量已經超過任何一個人所能消化的範圍，因此，如何幫助使用者節省時間在搜尋所需資訊上的議題漸漸受到重視。資訊篩選系統(filtering system)的產生可以改善資訊超載的情形。這類型篩選系統又可分為以內容導向式的資訊篩選系統(content-based filtering system)以及協同式資訊篩選系統(collaborative filtering system)兩大類。內容導向篩選系統可以收集並分析使用者在系統上的使用習慣及偏好，進而自動篩選出適合的資訊給使用者。另外，協同式篩選系統會根據有相同興趣的群體成員所擷取過的資訊作為推薦參考項目，而此方式的系統必須辨識相近的使用者，即找出最近鄰居，如何有效而快速地為使用者進行與鄰居的相似度辨識，是協同式篩選系統的首要課題。

在進行相似度的辨識上，協同式篩選系統必須收集所有使用者的相關使用資料，資料來源又可分為顯性回饋(explicit feedback)及隱性回饋(implicit feedback)資料兩種。顯性回饋資料為透過使用者明顯表達而建置的回饋資料，例如要求使用者輸入問卷資料、讓使用者對推薦項目進行評分等，此方式可藉由較少的資料量來取得直接而正確的使用者回饋訊息；隱性回饋資料則是由資訊系統收集使用者行為的各種資料再進行萃取，例如收集使用者的閱覽時間或使用者喜愛的文章等資料，再依據各種行為模式進行推論使用者可能的偏好。有鑑於顯性回饋較易取得使用者資料的特性，因此本研究採用顯性回饋來取得使用者資訊，並以此為依據當作使用者相似度辨識的基礎。

就受訪者而言，以個人口語(語意性措辭)表達對事物的態度，事實上是最直接且最容易，因此可運用於篩選系統中協助各使用者找尋偏好的近似鄰居，不僅能提高分群的準確率，更能有效縮短使用者興趣偏好比對上所耗費的時間。因此，為了更精確瞭解使用者資料，本研究將在協同式篩選系統上結合模糊語意分析法(fuzzy linguistic analysis)，以提供較細緻的最近鄰居挑選機制，並增加評估最近鄰居的準確度。

2. 文獻探討

本研究主要運用模糊語意分析法於協同式篩選系統中，以協助使用者找尋其最近鄰居。以下針對模糊語意分析法、協同式篩選系統及最近鄰居法三個部份，依序作相關的文獻探討。

2.1 模糊語意分析法

人類的思維及複雜的自然語言中都存在著不確定性，而且受訪者的認知型態也是不盡相同，在面對問卷的語意性措辭的感受上亦是模糊的，不適合用單一化的數值加以表達。因此，Zadeh(1973)認為運用模糊理論(fuzzy theory)操作不確定問題是相對較好的工具。而語意方法論(linguistic approach)與相關的模糊邏輯也經常以模糊理論為發展的基礎(Zadeh,1975)，受訪者以語意性措辭如：非常不同意、不同意、普通、同意與非常同意，對問題表達認知程度，而模糊語意(fuzzy linguistic)能更進一步衡量受訪者對該語意

性措辭，以語意區間評估感受上的強度。對不同的受訪者而言，個人認知的感受強度會有所不同，因此應用模糊理論以可重疊區間值的形式，具有連續性與不等長的語意區間，較符合實際狀況且能獲得較精確的結果。

在模糊語意的研究中，Bonissone(1980)曾提出如何在相關的語意中將語意措辭轉換成模糊集合，以及模糊數的演算方法。Chen and Hwang(1992)則提出一個應用模糊理論，將語意性措辭模糊數(fuzzy number)轉換為明確數值(crisp number)。模糊語意運用在李克特量表上的作法，稱之為模糊語意問卷(fuzzy linguistic scale)，其目的在於能完整地保留受訪者對於語意性措辭的真實想法與意見，而其優點在於語意區間的模糊資料可用語意性措辭來表示，降低模糊問卷填答時的困難度。

2.2 協同式篩選系統

由於網際網路及資訊數位化快速發展的影響，數位資訊量在網路上已無法估算，使用者往往得投入大量的時間去取得所需的資訊。因此，Resnick and Varian(1997)提出了協同式篩選技術以透過別人的篩選來產生決策，Herlocker, et al. (2000)認為協同式篩選系統的定義，是經由連接與個人具有相同興趣的使用者來預測使用者對一樣項目或資訊的喜好程度；Roh, et al.(2003)也解釋協同式篩選是一種知識分享的技術，Sarwar, et al.(1998)更認為篩選系統適合用統計及知識探索的技術以解決項目或資訊篩選的問題。協同式篩選的主要工作在使用者資料不足的情況下，能正確的推測其偏好或未知的興趣，並快速的篩選資訊給使用者。但由於現行的許多篩選系統其篩選的產生過程不夠透明，使用者無法得知篩選的依據。因此 Herlocker, et al.(2000)認為要讓使用者信服系統的篩選結果，則必須讓使用者看見篩選背後的推理機制。

良好的篩選系統需要精確的推理及預測，因此許多學者提出協同式篩選系統演算法，以改善使用者分群及預測篩選的資訊，例如 Lee, et al.(2003)考慮到應用在網站上的協同式篩選系統必須有效及自動的分類群體，並且要篩選高品質的資訊給具有相似興趣的使用者，因而建構出一個運用模糊理論(fuzzy)來產生篩選結果的協同式篩選系統；Li, et al.(2005)則以多樣性的興趣(multiple-interests)及多樣性的滿足(multiple-content)的觀點設計協同式篩選系統，以彌補協同篩選系統精確度不足的問題。然而，資訊篩選的準確度與回應的速度往往難以同時兼俱，為了改善此問題，Roh, et al.(2003)結合了協同式篩選演算法與機器學習方法，設計一個兼具品質與速度的篩選系統，結果確時增加準確度並減少運算時間。

協同式篩選在實務的運用上也有許多案例，如 Cohen and Fan(2000)將協同式篩選演算法結合網頁搜尋引擎，列出相關偏好的網站篩選給使用者；Hayes and Cunningham(2004)曾提出一個自動化的協同式篩選系統(automated collaborative filtering)，運用線上及時(real-time)的方式擷取使用者所喜愛的音樂類型，並篩選新的音樂給使用者；Lee, et al. (2005)也曾經使用迴歸模型於協同式系統上，分析使用者購物籃資料以達到更精確的分類及預測；Schafer, et al.(1999)認為協同式篩選能為電子商務網站帶來更多的商業價值，例如：Amazon、eBay 等等。Amazon 專注將協同式篩選系統運用於書籍與顧客等資料的分析，藉由書籍的購買人數資料加以分析，以推薦作者之最新

著作給特定顧客；eBay 則買賣雙方之評比與信用度資料，經篩選後推薦給其他買家作為參考。

2.3 最近鄰居

在所有的使用者資料中，找尋與目標個體最相近的一個，此即為最近鄰居的挑選工作。在最近鄰居的挑選方法上，常以歐基里德距離公式(1)計算出兩者間的距離(Ku, et al., 2005)，再找出兩兩間距離最短者做為目標個體的最近鄰居，歐式的距離 $D(x,y)$ 的計算為

$$D_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} , \tag{1}$$

n =空間維度大小，

x_i =使用者 x 於第 i 項問題的記錄值，

y_i =使用者 y 於第 i 項問題的記錄值。

最近鄰居的找尋常被運用在不同領域中，如知識發掘(knowledge discovery)、資料探勘(data mining)、圖像(pattern)的辨識及機器學習(machine learning)等議題(ARYA,1998)。最近鄰居的挑選速度常受限於資料的大小，當資料多且考慮挑選的準確性時，最近鄰居的挑選可能會耗費相當多的時間。因此，最近鄰居的研究大部份在改良其演算方法，以提高找尋最近鄰居的準確性與速度。在實務上如 Sarwar, et al.(2000)提出以最近鄰居演算法結合協同式篩選系統，改進在電子商務網站上資訊篩選的品質；Jensen, et al.(2003)也曾提出以最近鄰居演算法協助使用者，在公路網路(road networks)行進中可查詢目標位置的系統，本文則將以模糊語意法增加最近鄰居挑選的語意完整度。

3. 研究方法

本研究以「終極網頁寫作研究班」網頁製作之電子報網站為例，運用模糊語意分析法找尋目標使用者的「最近鄰居」。本文的研究架構主要分為三個主要階段，1.資料收集 2.離異度計算 3.找尋最近鄰居。相關步驟整理如圖 1 所示，並逐步說明如下。

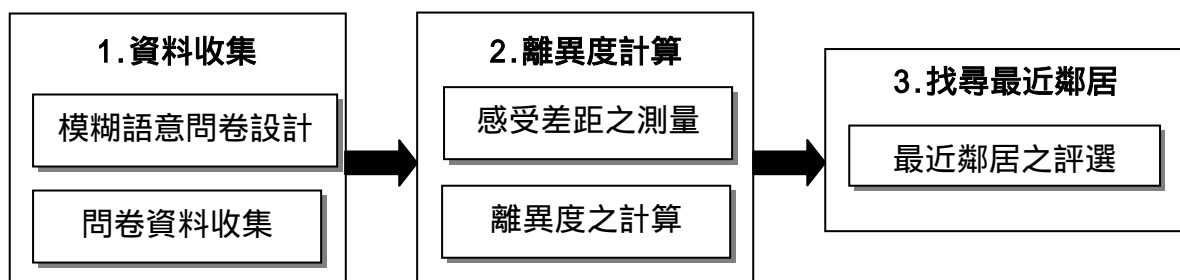


圖 1 研究流程圖

3.1 資料收集階段

資料收集階段主要的工作是收集使用者對於主題的涉入(involvement)程度資料，收

集的方法是以線上網路問卷的方式，在使用者初次進入篩選系統時即進行問卷調查，以收集各使用者對於主體的涉入資料。

在問卷設計主要的題目為針對資訊內容之涉入量表，再於每一涉入問題選項中加入語意測量區間量表，如表 1 所示。涉入量表係參考 Zaichkowsky(1994)的十題涉入問題，採李克特五點量表設計。由於本文是採取顯性回饋來收集使用者的資料，為降低使用者的拒答率，因此必需使問卷題項越少越好。本文採取 Zaichkowsky(1994)的涉入問卷主要的原因，不僅在於此量表在涉入議題的探討上已相當具代表性，同時其量表僅有十題，也可達到題數少的目標。在語意測量部份，則於每個涉入問題之五點量表下各有三個格位的區間分數，區間分數是一組介於 0 與 1 之間的模糊數，第一格之分數代表填答人對於該選項最低的認知強度，第三格分數則為填答人對該選項的最高認知強度，第二格分數則代表填答人對此選項的平均認知強度。例如，若有一位使用者在「非常不同意」這個選項下的區間分數填寫成(0,0.1,0.2)，則代表該使用人對於該「非常不同意」選項的認同強度只有 0 到 0.2 之間，而認知強度其中間值是 0.1，因此對於此選項的語意上，該使用人的認知是相對較弱的。

問卷設計完成後即將問卷置於電子資訊推薦系統中，並於每位新進使用者進入該系統時即填寫此一線上問卷，並將填寫後資料建置於資料庫中。

表 1 語意測量區間量表

1. 電子報提供網頁製作資訊對您而言是重要的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															
2. 電子報提供網頁製作資訊對您而言是相關的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															
3. 電子報提供網頁製作資訊對您而言是有義意的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															
4. 電子報提供網頁製作資訊對您而言是有價值的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															
5. 電子報提供網頁製作資訊對您而言是生活上必要的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															
6. 電子報提供網頁製作相關資訊對您而言是您感興趣的															
感受程度：	1.非常不同意			2.不同意			3.普通			4.同意			5.非常同意		
語意區間：															

表 1 語意測量區間量表(續)

7. 電子報提供網頁製作相關資訊對您而言是不可缺乏的														
感受程度：	1.非常不同意	2.不同意	3.普通	4.同意	5.非常同意									
語意區間：														
8. 電子報提供網頁製作資訊對您而言是心動的														
感受程度：	1.非常不同意	2.不同意	3.普通	4.同意	5.非常同意									
語意區間：														
9. 電子報提供網頁製作資訊對您而言是值得擁有的														
感受程度：	1.非常不同意	2.不同意	3.普通	4.同意	5.非常同意									
語意區間：														
10. 電子報提供網頁製作資訊對您而言是想要得到的														
感受程度：	1.非常不同意	2.不同意	3.普通	4.同意	5.非常同意									
語意區間：														

3.2 離異度計算階段

當有使用者進行資料查詢時，系統將進行最近鄰居的尋找。首先，系統將以各使用者之十題涉入問卷資料做為比較基礎，進行離異度的計算，此階段又分為感受差距測量與離異度計算兩個步驟。

3.2.1 感受差距之測量

首先要計算使用者兩兩之間對各問題感受的語意區間差距，即感受差距。感受差距是以各別使用者對特定問卷題項所選的語意區間，分別計算使用者與使用者間各問題之語意區間分數的差距。令使用者 x 與 y 對於問卷題項 i 的模糊區間分別為 (L_x, M_x, R_x) 與 (L_y, M_y, R_y) ，L、M 與 R 分別代表模糊區間之最低、平均與最高強度，則對於此題項 i，x 與 y 的感受差距 $fd_{i(x,y)}$ 為 $(L_{ix}, M_{ix}, R_{ix}) - (L_{iy}, M_{iy}, R_{iy}) = (L_{ix} - R_{iy}, M_{ix} - M_{iy}, R_{ix} - L_{iy}) = (u_{iL}, u_{iM}, \mu_{iR})$ ， u_{iL} 、 u_{iM} 與 μ_{iR} 分別為使用者 x 與 y 對於題項 i 感受差距之最低、平均與最高強度值。感受差距的計算結果仍是一項模糊數，無法明確的表示其之間的差距值。因此，必須再根據解模糊化方程式計算出明確的差距值 $fd^*_{i(x,y)}$ ：

$$fd^*_{i(x,y)} = |[(\mu_{iR} - \mu_L) + (\mu_{iM} - \mu_{iL})] / 3 + \mu_{iL}|, \tag{2}$$

例如：使用者 A 與使用者 B 對於問題 1 的語意區間分別為(0,0.1,0.2)與(0.8,0.9,1)，對兩位使用者而言，問題 1 的感受差距即是(0,0.1,0.2)-(0.8,0.9,1) =(-1,-0.8,-0.6)，再經由解模糊化計算， $fd^*_{1(A,B)} = |[(0.4)+(0.2)]/3-1| = 0.8$ 。所以，使用者 A 與使用者 B 對於問題 1 的感受差距值為 0.8。在計算出兩個使用者間的對於某題項的感受差距後，可依此方法，依次將兩兩使用者在其餘題項的感受差距計算出來。

3.2.2 離異度計算

兩個使用者的離異度計算，主要是計算兩者各題項間的意見差異，再加上感受差距做為權重，使用者 x 與 y 的離異度 $Sd^*_{(x,y)}$ 為

$$Sd^*_{(x,y)} = \sqrt{\sum_{i=1}^n fd^*_{i(x,y)} (r_{xi} - r_{yi})^2} \quad (3)$$

n =問卷題數，

r_{xi} =使用者 x 於第 i 項問題的記錄值，

r_{yi} =使用者 y 於第 i 項問題的記錄值，

$fd^*_{i(x,y)}$ =使用者 x 與 y 於第 i 項問題的感受差距值。

例如，表 2 中使用者 A 與 B 的離異度經計算後為 5.76，以此類推，可得到使用者 A 與其它使用人的離異度，離異度越大，代表兩個使用者的使用習慣越不相似。

表 2 離異度範列表

使用者 \ 問題	題 1	題 2	題 3	題 4	題 5	題 6	題 7	題 8	題 9	題 10	離異度
使用者 A	1	4	3	2	5	3	5	1	5	3	
使用者 B	5	4	5	4	5	1	5	5	3	2	
$fd^*_{(A,B)}$	0.8	0	0.43	0.5	0.03	0.4	0.03	0.83	0.4	0.2	5.76

3.3 找尋最近鄰居

依照上階段離異度計算，系統可逐步計算出目標使用者與其他使用者間的離異度，由於離異度與相似度呈反向關係，因此必須挑選離異度最小者做為目標使用者的最近鄰居。

4. 實証分析

本文以模糊語意尋求最近鄰居，相關的運用於「終極網頁寫作研究班」網頁製作之電子報網站上，以所屬之會員使用者為分析對象來進行比對分析的過程，從結果中找出各會員使用者的最近鄰居。

4.1 使用者資料收集

使用者資料的收集是經由電子報網站之會員使用者，以線上問卷的方式填答有關評估興趣偏好的模糊語意問卷。問卷經使用者填答確認後，記錄方式如表 3 所示。在表 3 中記錄使用者共收集有 50 人，編號從 U_1 到 U_{50} ，其中使用者 U_1 在問卷題項 1 選擇 1(非常不同意)，而其對題 1 之語意措辭區間模糊數為 0 至 0.2 之間，中間值為 0.1，因此在表 3 中表示為 1(0,0.1,0.2)，其餘依次類推。

表 3 使用者對各準則項目之語意區間模糊數記錄表

問題 使用者	題1	題2	題3	題4	題5
	U ₁	1(0,0.1,0.2)	4(0.7,0.8,0.9)	3(0.4,0.5,0.6)	2(0.2,0.3,0.4)
∴	∴	∴	∴	∴	∴
U ₅₀	3(0.3,0.4,0.5)	5(0.5,0.6,0.7)	1(0.8,0.9,1)	5(0.2,0.3,0.4)	3(0,0.1,0.2)
問題 使用者	題6	題7	題8	題9	題10
	U ₁	3(0.4,0.5,0.6)	4(0.8,0.9,1)	5(0,0.1,0.2)	2(0.8,0.9,1)
∴	∴	∴	∴	∴	∴
U ₅₀	2(0.2,0.3,0.4)	3(0.3,0.4,0.5)	5(0.8,0.9,1)	2(0.2,0.3,0.4)	1(0,0.1,0.2)

4.2 離異度計算

如前所述，離異度的計算首先必須計算出其它使用人與目標使用者的感受差距，再計算兩者的離異度。本文以編號 U₁ 的使用者為目標使用人，並以 U₂ 做為計算離異度的範例，表 4 首先計算使用者 U₁ 與 U₂ 問卷各題項的感受差距，其中以第一題的感受差距最大(0.73)，代表兩者對於本題的認知有所差異。在計算兩者之感受差距後，再配合兩位使用者 U₁ 與 U₂ 實際的選擇，進而求得兩者的離異度(4.34)，列如表 5。

表 4 U₁ 及 U₂ 使用者之感受差距值計算範例表

使用者 問題	U ₁	U ₂	感受模糊距離 fd _(U₁,U₂)	感受差距值 fd* _(U₁,U₂)
題1	(0,0.1,0.2)	(0.8,0.8,0.9)	(-0.9,-0.7,-0.6)	0.73
題2	(0.7,0.8,0.9)	(0.8,0.8,0.9)	(-0.2,0,0.1)	0.03
題3	(0.4,0.5,0.6)	(0.9,1,1)	(-0.6,-0.5,-0.3)	0.47
題4	(0.2,0.3,0.4)	(0,0,0.1)	(0.1,0.3,0.4)	0.27
題5	(0.8,0.9,1)	(0.7,0.7,0.8)	(0,0.2,0.3)	0.17
題6	(0.4,0.5,0.6)	(0,0,0.1)	(0.3,0.5,0.6)	0.47
題7	(0.8,0.9,1)	(0.2,0.2,0.3)	(0.5,0.7,0.8)	0.67
題8	(0,0.1,0.2)	(0,0,0.1)	(-0.1,0.1,0.2)	0.07
題9	(0.8,0.9,1)	(0.7,0.7,0.8)	(0,0.2,0.3)	0.17
題10	(0.4,0.5,0.6)	(0,0,0.1)	(0.3,0.5,0.6)	0.47

表 5 U₁ 及 U₂ 使用者之相異度計算範例表

問題 使用者	題 1	題 2	題 3	題 4	題 5	題 6	題 7	題 8	題 9	題 10	離異度 Sd* _(1,2)
使用者 U ₁	1	4	3	2	5	3	5	1	5	3	
使用者 U ₂	4	4	5	1	4	1	2	1	4	1	
fd* _(U₁,U₂)	0.73	0.03	0.47	0.27	0.17	0.47	0.67	0.07	0.17	0.47	4.34

4.3 近似鄰居之評選

依照上階段相似度的計算程序，資訊篩選系統可依序計算出目標使用者 U₁ 與其他使用者的離異度，如表 6 所示，其中表 6 之離異度值以由小至大排序。

表 6 使用者 U₁ 近似鄰居評選

使用者	離異度	使用者	離異度	使用者	離異度	使用者	離異度	使用者	離異度
U ₃₁	0.12	U ₁₉	1.86	U ₃₄	3.32	U ₂₀	4.6	U ₄₆	7.12
U ₂₄	0.34	U ₁₂	2.19	U ₂₇	3.35	U ₄₂	4.63	U ₃₇	7.28
U ₁₇	0.47	U ₃	2.29	U ₄₇	3.54	U ₂₈	5.14	U ₃₂	7.59
U ₄	0.5	U ₄₃	2.36	U ₅	3.76	U ₂₆	5.5	U ₂₅	8.27
U ₄₀	0.53	U ₂₃	2.4	U ₉	3.89	U ₁₃	5.76	U ₁₅	8.61
U ₁₈	1.12	U ₃₅	2.76	U ₁₆	3.91	U ₃₈	6.36	U ₇	8.84
U ₄₅	1.34	U ₂₂	2.77	U ₃₃	4.12	U ₄₁	6.75	U ₄₄	9.08
U ₄₉	1.4	U ₃₉	2.81	U ₁₀	4.17	U ₁₄	6.82	U ₄₈	9.23
U ₃₆	1.53	U ₃₀	3.2	U ₂	4.34	U ₁₁	7	U ₂₁	9.69
U ₅₀	1.6	U ₈	3.31	U ₆	4.57	U ₂₉	7.07		
平均數：4.27，標準差：2.691									

由於離異度與相似度呈反向關係，因此在本例中可挑選使用者 U₃₁ 做為 U₁ 的最近鄰居。另外，整體挑選結果的離異度平均值為 4.27，標準差為 2.691。

4.4 評選結果比較

若不考慮模糊語意之感受差距值，單粹只由五點量表的等距尺度來計算各使用者之間的離異度，則使用者 U₁ 與 U₂ 之間的離異度為 5.74(表 7)，大於具模糊語意之離異度(4.34，表 2)。

表 7 未考慮模糊語意之 U₁ 及 U₂ 使用者離異度計算範例表

問題 使用者	題 1	題 2	題 3	題 4	題 5	題 6	題 7	題 8	題 9	題 10	離異度
使用者 U ₁	1	4	3	2	5	3	5	1	5	3	
使用者 U ₂	4	4	5	1	4	1	2	1	4	1	
											5.74

依次類推，本研究將所有與使用者 U_1 的鄰居依此方法依序計算出與使用者 U_1 的離異度(表 8)，以做為和具模糊語意權重的比較基礎。表八資料結果顯示，使用者 U_1 的最近鄰居將評選為使用者 U_{19} (離異度 0.23)，與前面模糊語意的挑選結果已有不同。

表 8 使用者 U_1 最短距離鄰居評選

使用者	離異度	使用者	離異度	使用者	離異度	使用者	離異度	使用者	離異度
U_{19}	0.23	U_{17}	2.21	U_{39}	3.68	U_{20}	5.61	U_{46}	7.51
U_{24}	0.31	U_{12}	2.52	U_{22}	3.96	U_{26}	5.82	U_{15}	7.88
U_{31}	0.33	U_{35}	2.77	U_{47}	4.03	U_{28}	5.94	U_{48}	8.28
U_4	0.42	U_{43}	2.86	U_5	4.19	U_{42}	6.14	U_{25}	8.59
U_{45}	1.28	U_{23}	2.89	U_9	4.33	U_{13}	6.46	U_{37}	8.84
U_{36}	1.31	U_3	2.94	U_{10}	4.59	U_{38}	6.68	U_7	9.47
U_{40}	1.33	U_{27}	3.2	U_6	4.82	U_{41}	6.85	U_{21}	9.66
U_{49}	1.5	U_{34}	3.28	U_{16}	5.12	U_{11}	6.88	U_{32}	9.72
U_{18}	1.53	U_{30}	3.35	U_2	5.38	U_{14}	7.11	U_{44}	9.85
U_{50}	1.64	U_8	3.59	U_{33}	5.5	U_{29}	7.27		
平均數：4.69，標準差：2.788									

若進一步將以兩次(表 6 與表 8)的評選內容做比較，則具模糊語意的評選結果在離異值平均數(4.27)，比未具權重的方法(4.69)低(p 值=0.0001)。儘管如此，由於兩者方法皆受到使用者主觀認定所影響，因此在此並未能有充份證據顯示，模糊語意法在挑選最近鄰居可以較優於未具權重的方法。然而，相較於未具權重的方法，以模糊語意進行最近鄰居的尋找，確較能完整地保留受訪者對於語意性措辭的真實想法與意見，在操做上也較細緻且提供較能逼近於最近鄰居的機制，並可以獲致良好的挑選結果。

5. 結論

在網際網路快速的發展之下，資訊量已超出人力負荷的範圍，因此各種因應的資訊篩選系統因孕而生。本研究是以協同式篩選系統為基礎，配合模糊語意法的運用，嘗試找出目標使用者的最近鄰居。

本研究在尋求最近鄰居的做法上，共分成三大步驟，首先在篩選系統設置資訊涉入問卷，並要求每位新進使用人進行填寫，以收集所有使用者對於系統提供資訊的涉入狀況。一旦有目標使用人欲進行最近鄰居的挑選時，系統首先以模糊語意法找出目標使用人與鄰居間的感受，再進行兩者間的離異度計算，最後並進行最近鄰居的評選。

本研究並以「終極網頁寫作研究班」網頁製作之電子報網站的 50 位使用者進行實證，並順利找出最似鄰居。資料分析結果並進一步與未具權重的挑選方法做比較，資料分析結果顯示，模糊語意法可以在近似鄰居的找尋上獲得良好的結果。在後續的發展部份，由於本研究目前僅以使用者的涉入問卷資料做為最近鄰居的挑選基礎，未來可嘗試增加其它分析構面並收集更多的使用資料，以更精確挑選出真正的最近鄰居。

參考文獻

1. ARYA, S., D. M. Mount, N. S. Netanyahu, R. Silverman and A. Wu (1998), "An Optimal Algorithm for Approximate Nearestneighbor Searching in Fixed Dimensions," *Journal of the ACM*, 45, pp.891-923.
2. Bonissone, P. P. (1980), "A Fuzzy Sets Based Linguistic Approach : Theory and Applications", *Proceedings of the 12th conference on Winter simulation* , pp.99-111.
3. Chen, S. J. and C. L. Hwang (1992), *Fuzzy multiple attribute decision making method and application*, NewYork: Springer-Verlag.
4. Cohen, W. W. and W. Fan (2000), "Web-collaborative Filtering: Recommending Music by Crawling the Web," *Computer Networks*, 33(1/6), pp.685-698.
5. Hayes, C. and P. Cunningham (2004), "Context Boosting Collaborative Recommendations," *Knowledge- Based Systems*, 17(2/4), pp.131-138.
6. Herlocker, J. L., J. A. Konstan and J. Riedl (2000), "Explaining Collaborative Filtering Recommendations", *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pp.241-250.
7. Jensen, C. S., J. Kolárřvr, T. B. Pedersen and I. Timko (2003), "Nearest Neighbor Queries in Road Networks," *Proceedings of the 11th ACM international symposium on Advances in geographic information systems*, pp.1-8.
8. Ku, W. S., R. Zimmermann, H. Wang and C. N. Wan (2005), "Adaptive Nearest Neighbor Queries in Travel Time Networks," *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, pp.210-219.
9. Lee, D. S., G. Y. Kim and H. I. Choi (2003), "A Web-based Collaborative Filtering System," *Pattern Recognition*, 36(2), pp.519-526.
10. Lee, J. S., C. H. Jun, J. Lee and S. Kim (2005), "Classification-based Collaborative Filtering Using Market Basket Data," *Expert Systems with Applications*, 29(3), pp.700-704 .
11. Li, L., L. Liu and X. Li (2005), "A Hybrid Collaborative Filtering Method for Multiple-interests and Multiple-content Recommendation in E-Commerce," *Expert Systems with Applications*, 28(1), pp.67-77.
12. Resnick, P. and H. R. Varian (1997), "Recommender System", *Communications of the ACM*, 40(3), pp.56-58.
13. Roh, T. H., K. J. Oh and I. Han (2003), "The Collaborative Filtering Recommendation Based on SOM Cluster-indexing CBR," *Expert Systems with Applications*, 25(3), pp.413-423.
14. Sarwar, B. M., J. A. Konstan, A. Borchers, J. Herlocker, B. Miller and J. Riedl (1998), "Using Filtering Agents to Improve Prediction Quality in The Group Lens Research Collaborative Filtering System," *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, pp.345-354.

15. Sarwar, B., G. Karypis, J. Konstan and J. Riedl (2000), "Analysis of Recommendation Algorithms for E-commerce," *Proceedings of the 2nd ACM conference on Electronic commerce*, pp.158-167.
16. Schafer, J. B., J. Konstan and J. Riedl (1999), "Recommender Systems in E-commerce," *Proceedings of the 1st ACM conference on Electronic commerce*, pp.158-166.
17. Zadeh, L. A. (1973), "Outline of A New Approach to The Analysis of Complex Systems and Decision Process," *IEEE Transactions on Systems, SMC-3*, pp.28-44.
18. Zadeh, L. A. (1975), "The Concept of A Linguistic Variable and Its Application to Approximate Reasoning," *Information Sciences*, pp.199-249.
19. Zaichkowsky, J. L. (1994), "The Personal Involvement: Reduction, Revision, and Application to Advertising", *Journal of Advertising*, 23, pp.59-70.