

Optimal Assembly of Test construction by

Ant Colony Optimization

以螞蟻演算法進行測驗卷編製

何祖鳳

銘傳大學資訊工程學系

tfho@mcu.edu.tw

高嘉陽

銘傳大學資訊工程學系

s5366055@ss24.mcu.edu.tw

徐熊健

臺北市立教育大學資訊科學系

sjshyu@tmue.edu.tw

Abstract

摘要

隨者資訊科技的蓬勃發展，以電腦題庫產生測驗卷的作法，已日漸受到重視。早期，從題庫中挑選考題的方法，大多採用隨機選取，此法雖然簡便，卻無法滿足教師更多的需求；因此，為了使測驗卷能達到測試的效果，將題庫中的題目加上鑑別度、關聯度、難易度、猜測度、作答時間等屬性，以試題反應理論為本研究之理論基礎，根據教師特定的需求下，電腦自動由題庫中挑選符合條件的題目，產生一份最佳測驗卷。目前，針對測驗編製的問題，已有研究學者應用經驗演算法（Heuristic Algorithm），以快速產生一份很好的測驗卷。本研究以螞蟻族群演算法自大量題庫產生測驗卷，經由實驗得知，在小數量的情況下，螞蟻族群演算法比其它經驗演算法取得的解稍好，在題庫數量越來越大時，其他的經驗演算法已無法取得解時，而螞蟻演算法仍可以找出解，且螞蟻較使用更少的時間來完成一份試卷。

關鍵詞：題庫、測驗卷、螞蟻族群演算法，經驗法則，試題反應理論

Generating test-sheet from item banks has attracted much attention in the area of computer assisted learning. Most methods of selecting test item from item banks so far are based upon random selection. Although it is easy but it is hard to satisfy the different needs of teachers. Further, every test item of item bank affects the degree of discrimination, degree of association, degree of difficult, degree of guess, length of test time, and so on. In this paper, we adopt the item response theory to formulate the test-sheet composition. According to the specific demand from a teacher, our proposed approach generates a satisfactory test-sheet satisfying all criteria from a certain item bank. Our approach is based upon Ant Colony Optimization. According to the experimental results, the proposed ACO algorithm is comparable to the 0/1 linear programming approach, and it is more efficient. The ACO algorithm is still efficient when the problem size becomes large.

Keyword : item bank, test-sheet, Ant Colony Optimization Algorithm, Heuristic Algorithm, item response theory

一、簡介

近年來，電腦技術持續的發展，提升了電腦的運算速度，眾多的研究結果擴展電腦相關的應用發展。電腦能做的事情越來越複雜，意味者電腦本身能夠執行一定程度有智慧的行為。網際網路的技術實現了遠距教學，讓知識傳授的方式已不再侷限於傳統教室的黑板上，透過電腦平台結合網際網路，學習行為發生在任何場所任何時間內。在這樣的趨勢中，透過電腦來實行測驗的方式也就越來越多。如 GRE(Graduate Record Examinations)為 GRE 委員會委託美國教育測驗服務社舉辦的世界性測驗。GRE 在 1993 年開始實施部份以電腦為媒介的測驗方式，直至 1999 年，放棄以紙筆的傳統受測方式。

測驗內容是從整個教學過程中，擷取重要的內容知識化為選擇、問答等等的測驗題目，藉此測驗卷可以幫助老師憑斷學生的學習成果，也可以透過作答的情形了解學生的學習狀態。今天老師要花費不少的時間精力依照學生的程度、所學的內容，製作出一份難易適中，且能夠判斷出學生的學習成果的測驗卷。如果同時要製作多份試卷的話，那時間將會白白的浪費掉。如今在電腦能力不斷的提升下，本研究將製作試卷的工作轉移給電腦處理，冀望在電腦的幫助之下能更快速完成製作試卷的工作。

在測驗系統中，整份測驗卷的好壞由測驗內容所影響，難易度、辨別度等等表示為測驗題目的質量。由題庫挑選的題目不符合學生的程度，根據試題反應理論所提供的數學模式，描述學生能力與作答之間的關係，挑出對於不同程度的學生最合適的測驗卷。另一方面能影響好壞的關鍵就在於組成測驗卷的方式，一定的限制下，選擇符合這些限制的題目組成一份測驗卷，是一個 NP-Hard 問題 [13][18][21]。題庫至少有一定的數量，且題庫的數量會逐漸增加，這樣的情況下，組成測驗卷將變的更加複雜。

二、試題反應理論

試題反應理論有兩個基本概念[1,6][20]，受測者在某一試題中的表現情況，可以用一組因素來加以預測或解釋，這組因素稱為潛在特質或能力；描述受測者的能力高低與某一個測驗項目的表現情形，可以用一組函數來加以解釋，此函數稱作試題特徵函數或試題特徵曲線(item characteristic curve)，即能力不同的受測者得分點所構成的曲線，就是該測驗試題的試題特徵曲線，其意義在於受測者的能力決定了他在此題答對的機率，當受測者的能力越高，他在這題的答對的機率也就越高，而將所有的試題特徵曲線加總起來，即便構成試卷特徵曲線(test characteristic curve)；試題反應理論必須滿足下列四項基本假設，當這些假設都成立時，試題反應模式也才具有意義，如下所示：

1. 單向度(Unidimensionality)，受測者在作答的情況下，都是受到同一種能力或是潛在特質的影響，但是在實際的作答情形中，受測者可能受到許多因素影響如成就動機、考試焦慮、應試技巧、及其他人格特質等等，試題反應理論認為該測驗有一個能夠影響測驗結果的主要成份或因素，即符合單向度假設的基本要求。
2. 局部獨立性(Local independence)，受測者在對任何試題上的反應是獨立的，試題反應理論的能力因素，才是唯一能夠影響受測者在測驗試題上做反應的因素，當單向度基本假設成立時，局部獨立性假設也就成立。
3. 非速度測驗(Nonspeedness)，受測者的考試成績不理想，是因為其能力不足所導致，並非由於時間不夠答完所有試題所影響。
4. 知道—正確假設(Know--correct assumption) 考生知道試題的正確答案，必然會答對該試題，相反的，如果他答錯某一試題，則表示他不知道該試題的答案，因為人為的疏忽而導致成績不理想，不在本假設所考慮的範圍內。

試題特徵曲線是用來描述能力與其在試題上答對機率之間的一種數學關係，因此，根據參數的

多寡會有不同的試題反應模式，試題參數是指試題的難易度、鑑別度、猜測率等等用來描述試題特性的指標，常用的試題反應模式可分為以下三種

1. 一個參數對數形模式

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}} \quad (1)$$

2. 兩個參數對數形模式

$$P_i(\theta) = \frac{e^{a_i(\theta-b_i)}}{1+e^{a_i(\theta-b_i)}} \quad (2)$$

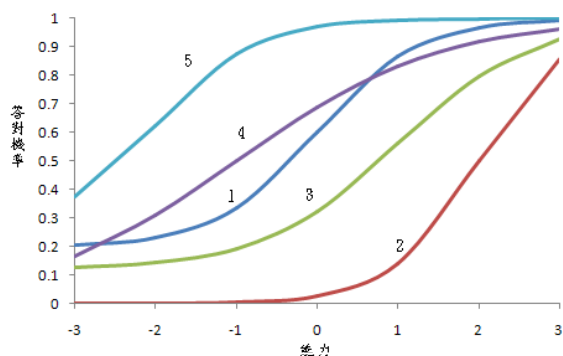
3. 三個參數對數形模式

$$P_i(\theta) = C_i + (1-C_i) \frac{e^{a_i(\theta-b_i)}}{1+e^{a_i(\theta-b_i)}} \quad (3)$$

$P_i(\theta)$ 表示能力為 θ 的學生答對 i 題的機率大小， a_i 表示題目 i 的鑑別度， b_i 為題目 i 的難易度， C_i 為題目 i 的猜測度， e 為 2.718 的指數， $P_i(\theta)$ 是介於 0 與 1 的一條曲線，以表一的試題參數為例，根據三個參數對數形模式畫出的試題特徵曲線，如圖一所示：

表一 五個試題的試題參數

測驗題目	參數		
	鑑別度	難易度	猜測度
1	1.6	0	0.2
2	1.8	2	0
3	1.2	1	0.12
4	0.8	-1	0
5	1.8	-2	0.25



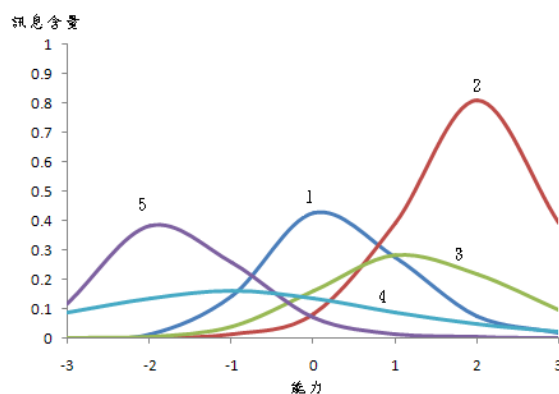
圖一 三個參數試題特徵曲線

試題反應理論應用到測驗編製上，最重要的一項工具就是訊息函數(information function)，利用試題訊息函數作為建立、分析、診斷測驗的依據，以三個參數對數形模式為例，試題訊息函數的定義如下：

$$I_i(\theta) = \frac{a_i^2(1-C_i)}{[C_i + e^{a_i(\theta-b_i)}][1 + e^{-a_i(\theta-b_i)}]^2} \quad (4)$$

$I_i(\theta)$ ：測驗題 i 在能力為 θ 上所提供的訊息量。

依照選擇的試題反應模式分析資料以後，可以得到該試題的訊息函數值，試題訊息函數在測驗的編製上，以及試題好壞的診斷上，佔有很重要的成份，因為它能表示試題對能力的訊息量大小，根據表一的資料所畫出的訊息函數如圖二所示：



圖二 表一中五個試題訊息函數

一份測驗卷的訊息函數是在 θ 值上的所有試題訊息函數之總和為 $I(\theta)$ ，根據訊息量大小可以得知該測驗適用於何種能力估計值範圍的測量，希望能力的估計值越精確，估計的標準誤 $SE(\hat{\theta})$ 越小越好，也就是測驗卷的訊息函數值要越高越好的，估計標準誤(standard error of measurement)如下所示：

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \quad (5)$$

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}} \quad (6)$$

三、選題策略介紹

好的方法不僅幫助我們挑到好的試題，也能加快測驗卷的編製。組合測驗卷問題中，已有許多研究學者提出各自的方法應用在此問題上。施測者以受測者的能力訂出目標訊息量，由題庫挑選出試題為一份測驗卷，目的在求目標訊息量與測驗卷訊息量的誤差越小越好。

1999年，孫光天等人以貪婪演算法作為選題的方法[8]，選題的過程可分為兩階段，首先在挑選試題時，計算目標訊息量與測驗卷訊息量之間的誤差，選擇誤差最小的試題加入測驗卷，挑選的題目數量到達測驗卷的限制時，接者不斷的從題庫挑選試題取代目前測驗卷的題目，促使該測驗能夠找到最佳解；同年孫光天等人以基因演算法進行選題[9]，利用基因交配、突變與複製等特性，經過演化找出問題的解集中，最佳解或近似最佳解，提供此測驗卷編製的方法給教育工作者一個有效的工具。

2004年，陳健一將基因演算法，加入貪婪演算法的取其最優的概念[12]，取代基因演算法亂數選取的策略，應用在選題方面上，實驗的結果比基因演算法或貪婪演算法都來的要好，因為貪婪演算法的概念使得演算法在時間的收斂能夠變得更快。

2002年黃家輝以灰色理論為基礎[11]，編製的過程分為初選與置換，初選方式有目標值平均訊息法及隨機選取法，置換方式有灰關聯試題置換法、灰關聯權重置換法、參數修正置換法和直接置換法。兩階段方法互相搭配成數種選題方式發展出一套灰色選題系統。

以上四種選題策略所使用的題庫數量皆為320題，屬於小數量的實驗，在挑選的時候只有考慮到題數的限制，因此另有學者提出更多的限制以符合現實生活中的實際需求，以下介紹在多個限制下的選題方式[5]。

1991年，van der Linden 等人，以線性規劃求解測驗編製的問題[13]，製作測驗卷同時考慮到試題內容的意義，作為挑選試題的參考依據，可以避

開單純以試題的好壞作為挑選的方法，造成測驗的內容與教師實際想測的內容有相當的差距。

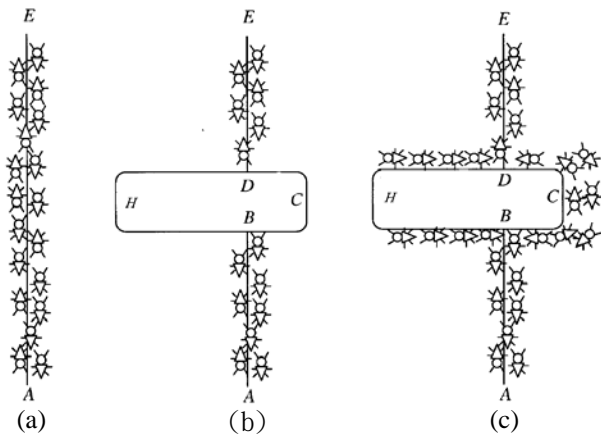
2003年巫眇鸞利用線性規劃結合選題策略[7]，先運用線性規畫鬆弛法求出實數解後，再結合分支界限法發展六種啟發式方法：鬆弛法、進位法、小數解最佳化、基底解最佳化、初現整數解與非0解最佳化。利用這些啟發式方法以求出編製測驗卷問題的最佳解，增加測驗之效率與正確性。

2003年孫光天以簡單的基因演算法為基礎，修改演化過程中所使用的參數及目標函數，為進階基因演算法[10]，應用在大量限制的選題問題上，能夠找出更好的解，提供實際的應用，使測驗卷的內容更符合設計者的需求。

以上三種方法，在測驗卷的題數之外，還考慮了測驗的內容，使得測驗卷的內容能夠更符合受測者的目標，但在實驗上題庫數量都1000題以下。因此當題庫數量變大時，以上方法所找的解是否能維持一定的品質和效率，乃一大挑戰。

四、螞蟻族群演算法介紹

螞蟻演算法最先由 Dorigo 和他的研究同伴提出來[15,17]，用來解決困難的組合最佳化的問題。模擬螞蟻在覓食過程中，她們在聚居地前往食物之間，會在經過的路徑上灑落費洛蒙，螞蟻可以偵測到地上所遺留的費洛蒙，灑落在地面上的費洛蒙會形成一條路徑，而此路徑可以指引螞蟻回到聚居地或者是到食物來源。當下一隻螞蟻在路徑的選擇上，則會有較大的偏好選擇費洛蒙較高的路徑。地上的費洛蒙透過螞蟻的行走而累積，同時費洛蒙也會以一定的速度蒸發於空氣中，隨者時間過去，較短的路徑，由於行走的時間較短故費洛蒙的累積速度較快，而路徑較長的費洛蒙，累積的量會比較慢，費洛蒙的消長之間，最後大部分的螞蟻會只走路徑短的。每隻螞蟻透過灑落費洛蒙在路徑上，分享自己的移動經驗給下一隻螞蟻作為參考，這種找尋路徑的過程可以視為一種透過合作的方式達到知識分享[22,23]。



圖三 螞蟻行徑方向的範例

以圖三為例子，A 點為螞蟻的家，E 點為食物處，在圖三(a) 中表示螞蟻在家與食物之間不斷的來回，此時如果在路徑中突然放入一個障礙物如圖三(b)，由 A 點出發的螞蟻到達 B 點時會遇障礙物，此時會面臨往左邊走或者往右邊走的選擇，同樣的由 E 點回來的螞蟻在 D 點也面臨相通的抉擇，在一開始螞蟻將以相同的機率選擇右邊或選擇左邊。由於路徑 BCD 的長度小於路徑 BHD，所以選擇路徑 BCD 的螞蟻比選擇路徑 BHD 的螞蟻更快走到點 E 如圖三(c)，因此在相同時間內，路徑 BCD 上的費洛蒙累積量將會大於路徑 BHD 的費洛蒙累積量。經過一段時間，其後的螞蟻在面對障礙物時，再選擇路徑上會有更大的偏好選擇路徑短的方向，最後所有螞蟻將會沿者最短的路走。

螞蟻族群演算法模擬螞蟻覓食的行為用以解決組合最佳化的問題，模仿螞蟻找尋食物的過程套用在問題求解的過程上，根據每次挑選的節點的有無來增減費洛蒙的量，在一定的時間後，費洛蒙量大的節點會形成一條路徑，此路徑即我們在此問題上所求得的答案。目前螞蟻族群演算法應用在最短路徑、旅行推銷員問題、資源分配問題、二次分派問題等組合最佳化問題上，已有非常良好的成果 [16]。

五、螞蟻族群演算法應用在測驗編製的問題

教師在進行測驗編製時，除了依照考生的能力，還會考慮測驗的內容有沒有包含不同類型的試題，或是作答時間必須在一定的範圍內等等需求，從題庫中挑選試題製成一份測驗卷。假設題庫中題目數量為 n ，每一個試題包含了鑑別度 a_i 、難易度 b_i 、猜測度 c_i ，使用的試題反應模式為三個參數對數形模式分析題庫資料，計算出題庫中的試題對於每個能力所提供的訊息量；老師出一份測驗卷，測驗的對象是程度高的學生，或者給所有程度的學生，根據學生的能力訂出目標訊息函數。老師在進行測驗時，如果還有其他的需求，像是測驗卷的作答時間要在一定的範圍，某一種類型的問題要挑選一定的題數等等限制。從題庫中挑選適當的題目，在符合所要求的限制之下，編製成一份測驗卷。測驗卷中的試題訊息函數與老師定的目標訊息函數的誤差能夠越小越好，即表示這份測驗卷越適合老師要測驗的對象，目標函式與限制式如下所示：

目標函式：

$$\min \sum_{k=1}^s (d_k - O_k)^2 \quad (7)$$

$$O_k = \sum_{i=1}^n I_i(\theta_k) x_i, \quad k = 1, 2, \dots, s \quad (8)$$

限制式：

$$\sum_{i=1}^n x_i r_{ij} \leq b_j, \quad j = 1, 2, \dots, m \quad (9)$$

$$x_i \in \{0, 1\} \quad i = 1, 2, \dots, n \quad (10)$$

n ：為題庫的大小；

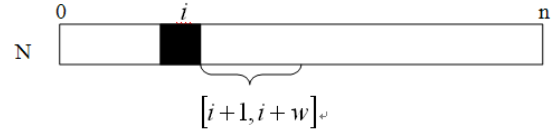
m ：問題的限制數；

s ：學生可依能力分為數群；

d_k ：老師所設定的目標函數；

O_k : 測驗卷中能力 θ_k 的訊息量總和；
 $I_i(\theta_k)$: 試題 i 對於能力 θ_k 所提供的訊息量；
 b_j : 問題的第 j 項的限制；
 r_{ij} : 試題 i 對 j 項限制的意圖；
 x_i : 表示測驗題目 i 是否被挑選，0 表示沒有挑到，1 則表示有被挑到。

$$\eta_i = \left[\sum_{k=1}^s I_i(\theta_k) \right] / s, \quad i = 1, 2, \dots, n \quad (11)$$



圖四 螞蟻挑選範圍圖示

(一)問題轉移

今天要從題庫中取出若干題，組合成一份適宜的測驗卷，首先初始化每個測驗题目的費洛蒙和計算出發發值；一開始先將螞蟻放在題庫上的任一點，螞蟻在移動的過程中，根據費洛蒙的強度與經驗法則的函式所計算出的偏好量，去選擇下一步該怎麼移動，利用一段記憶體的空間，記憶螞蟻選過的題目稱為已選名單；當所有螞蟻都挑選了初步解後，檢查螞蟻所找的解是否滿足所有的條件限制，如果有不滿足的則將他修復至符合所有限制，接者使用區域搜尋法更進一步改善解的結果，讓解的結果更接近最佳解，找出這回合組合測驗卷目標函數最好的稱之為回合最佳解，如果回合最佳解有大於全域最佳解的情況發生，則回合最佳解取代全域最佳解，最後在將費洛蒙更新過，以上步驟稱為一個回合。

(二)狀態轉移法則

螞蟻在移動的時候，會根據費洛蒙和啟發值去挑選下一步，題目被人挑選的情況反應在費洛蒙的強度上，費洛蒙越高者表示該題目較常被螞蟻選到，顯示出所有螞蟻對於此題目整體的經驗累積， τ_j 表示題庫中第 j 題的費洛蒙強度。當螞蟻在進行挑選時，會考慮一個適當的搜尋範圍大小為 w ，假設螞蟻 a_r 已選擇測驗項目 i ，下一步螞蟻挑選的範圍將由 $i+1$ 至 $i+w$ ，這樣的範圍稱為一個視窗 W 如圖四；而根據每個題目上的難易度、鑑別度、參測度等等相關訊息，透過訊息函數的公式(4)計算出每個能力值 θ 訊息函數的貢獻量，將這些訊息透過公式為這個測驗題目求出一個值， η_i 表示第 i 題的啟發值，公式如下所示：

$$p_i = \begin{cases} 1 & \text{if } q \leq q_0, \max \{ \tau_i \eta_i^\beta \}, i \notin tabu_r \\ 1/w & \text{if } q_0 \leq q \leq q_1, i \notin tabu_r \\ 0 & \text{otherwise } (q > q_1) \end{cases} \quad (12)$$

q_0, q_1 表示之間所佔的比例， $\beta (\beta \geq 0)$ 表示費洛蒙與啟發值之間何者比較重要，螞蟻 a_r 挑選第 i 題的機率如下公式所示，螞蟻會先隨機產生一個數字 $q (0 \leq q \leq 1)$ ，若 $q \leq q_0$ ，螞蟻會挑選視窗中 $\tau_i \eta_i^\beta$ 最大值得測驗題目，挑選過的題目則會放入 $tabu_r$ ；若 $q_0 \leq q \leq q_1$ ，螞蟻則會隨機挑選視窗任一題；若 $q > q_1$ ，表示不挑選這個視窗的任何一題，視窗會在往下移動 w 個距離，在產生一個隨機數字 $q (0 \leq q \leq 1)$ ，接著進行上述的動作。由公式(5)可知當螞蟻在挑選下一個題目時，其偏好較高的費洛蒙與啟發值。螞蟻在搜尋路徑時，會挑選目前可得的解，當可以挑選的解增加的越來越多，此時適度的將搜尋範圍縮小，就可以避免因大量的計算而拖累了時間。

(三)修復運算

當每隻螞蟻完成了初步的挑選，螞蟻會檢查所挑的解是否有滿足所有的條件限制，如果全部限制都滿足了稱之可行解，跳過此步驟，若有某一限制未達成則使用修復運算將不可行解改至可行解；挑選螞蟻 a_r 解中測驗題目 i_{drop} 與題庫未被挑選到的題目 i_{add} ，如果將 i_{drop} 與 i_{add} 交換可以更進一步的符合條件限制，則做交換的動作，修復的動作直到螞蟻 a_r 的解成為可行解即停止。

(四)區域搜尋法

在一回合中，我們找到了該回合的最佳解後，即使用區域搜尋法企圖改善這個回合最佳解，以期找到更好的全域最佳解。在此區域搜尋法是一個貪婪演算法，將此回合最佳解中的題目，一個一個去跟題庫中沒有被選到的題目去交換，試者在交換後能去得一個更好的解集合。

(五)費洛蒙更新法則

在一回合結束之時，我們會再目前為止最好的解組合上執行全域費洛蒙更新， T^+ 為目前為止最好的解組合， Q^+ 為測驗卷訊息函數總和最小之值，全域費洛蒙更新法則如下所示：

$$\tau_i = (1 - \rho) \times \tau_i + \rho \times \Delta \tau_i$$

$$\Delta \tau_i \begin{cases} \frac{Q^+}{n^{-1}} & \text{if } \langle i \rangle \in T^+; \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

$\rho(0 < \rho \leq 1)$ 表示費洛蒙蒸發的速率，所有題目的費洛蒙都會根據這個蒸發速率遞減，但是在最佳解集合中的每個題目，都會給予適當的獎勵，表示這些題目都是值得去選的，以便下一回合螞蟻在搜尋時，有多一點的機率再去選到這些好的題目。

(六)停止條件

當螞蟻演算法已經執行回合數的最大次數，或者是 CPU 已經執行一段時間，或者是經過數個回合仍然找不到更好的解可以取代目前最好的解，發生以上任何一個情形則螞蟻就會停止所有的動作，在此設定找到全域最佳解後，螞蟻持續三十回合仍然找不到更好的回合最佳解來更新全域最佳解，我們認為螞蟻在跑下去找到更好的解的機率不大，因此就結束搜尋的動作，或是螞蟻搜尋超過 200 回合尚未找到解答，故結束搜尋。

六、實驗結果與分析

實驗部分，依照 1991 年 van der Linden 等人所定義之問題，並將題庫的數量放大，求解並對照之間的差異。問題的目標函式定為這份測驗卷中的測驗卷訊息量越大越好，為避免集中某一能力值的訊

息量大，而其他能力值的訊息量小，故定測驗卷中訊息量中最少的為目標函式。

以下兩個實驗，第一個實驗為在同時製作兩份測驗卷。第二個實驗為在許多的條件限制下，來進行測驗卷的編製，測試的平台為個人電腦，CPU 為 Pentium III 1.13GHz 和 1GB 的記憶體，以 Matlab 與 C++ 程式語言開發。

1. 實驗一

題庫的內容為模擬資料，題庫包含三百個測試題目，每一個測驗題目均符合試題反應模式，鑑別度 $a_i = 0$ ，猜測度 $C_i = 0$ ，題目的難易度以標準常態分配的分布資料；本實驗是從題庫中各挑選測驗題目編製成兩份測驗卷，相同的測驗內容要出給不同的班級作為測驗，且兩份測驗卷不會有重複的情形發生，第一份挑過的題目第二份是不會在選的，而每一份測驗卷的測驗題目限制在 30 題。公式如下：

目標函式：

$$\max z \quad (14)$$

限制式：

$$\sum_{i=1}^{300} I_i(\theta_k) x_{it} - z \geq 0, \quad (15)$$

$$k = 1, 2, 3, \quad t = 1, 2,$$

$$\sum_{i=1}^{300} x_{it} = 30, \quad t = 1, 2, \quad (16)$$

$$x_{i1} + x_{i2} \leq 1, \quad i = 1, 2, \dots, 300, \quad (17)$$

$$x_{it} \in \{0, 1\}, \quad i = 1, 2, \dots, 300, \quad t = 1, 2, \quad (18)$$

$$z \geq 0. \quad (19)$$

表二 實驗一的結果

	試卷	試卷訊息函數			執行時間 (秒)
		I(-1)	I(0)	I(1)	
0/1 LP	A 卷	5.9106	7.4430	5.8534	13
	B 卷	6.0259	7.4533	5.7441	
ACO	A 卷	5.9724	7.4520	5.7968	0.125
	B 卷	5.9683	7.4456	5.7974	

表三 目標函式 z 值

	0/1 LP	ACO
z	5.7441	5.7968

實驗的結果顯示，由表三得知 0/1 線性規劃與螞蟻族群演算法所找的解相差不遠，螞蟻族群演算法的結果比 0/1 線性規劃稍微好了一點，且從表二的執行時間看來，螞蟻族群演算法明顯的快了 0/1 線性規劃。

2. 實驗二

題庫的內容為英文測驗的模擬資料，題庫包含六百個測試題目，每一個測驗題目均符合試題反應模式，鑑別度 $a_i = 0$ ，猜測度 $C_i = 0$ ，題目的難易度以標準常態分配的分佈資料，每一個測試題目還有一個時間假設將此題目做完需花多少時間，此部分資料呈均勻分配介於 [20,60]；題庫平均分成，單字、文法與閱讀測驗三個部分，每個共有 200 題的測驗題目，且每一個部份的前 100 題題目為多選題，其他則為填充題。現在施測者在編製測驗時考慮到幾項條件：

- I. 受測的對象包括不同程度的學生，因此測驗卷中每個能力值的訊息函數都應該相當接近。
- II. 測驗卷要包含 14 個單字，16 個文法和 10 閱讀測驗的題目。
- III. 要有 16 題的多選題，24 題為填充題。
- IV. 整份測驗卷的作答時間應該控制在 1500 秒以下。

目標函式：

$$\max z \quad (20)$$

限制式：

$$\sum_{i=1}^{600} I_i(\theta_k) x_i - z \geq 0, \quad k = 1, 2, 3, \quad (21)$$

$$\sum_{i=1}^{200} x_i = 14, \quad (22)$$

$$\sum_{i=201}^{400} x_i = 16, \quad (23)$$

$$\sum_{i=401}^{600} x_i = 10, \quad (24)$$

$$\sum_{i=1}^{100} x_i + \sum_{i=201}^{300} x_i + \sum_{i=401}^{500} x_i = 16, \quad (25)$$

$$\sum_{i=101}^{200} x_i + \sum_{i=301}^{400} x_i + \sum_{i=501}^{600} x_i = 24, \quad (26)$$

$$\sum_{i=1}^{600} x_i t_i \leq 1500, \quad (27)$$

$$x_i \in \{0,1\}, \quad i = 1, 2, \dots, 600, \quad (28)$$

$$z \geq 0. \quad (29)$$

表四 實驗二的實驗結果

題數	z		計算時間(秒)	
	ACS	0/1 LP	ACS	0/1 LP
600	7.84841	7.84631	4.6955	3.954
1200	7.85900	7.85223	5.7602	36.406
3000	7.86151	7.86157	20.4648	136.313
6000	7.86427	x	50.9257	x
9000	7.86438	x	84.2320	x
12000	7.86443	x	132.7660	x
15000	7.86444	x	157.0700	x
18000	7.86446	x	184.4870	x

在實驗結果中可分為兩部份，題庫數量小於 3000 的情況下 0/1 線性規劃所找的解與螞蟻演算法解，螞蟻演算法所找的解略好於 0/1 線性規劃，但在執行的時間上，螞蟻演算法明顯的快出 0/1 線性規劃的執行時間，且仔細檢視螞蟻所找到的解與 0/1 線性規劃找到的解，其實差距不大，而題數達 6000 題以上時，0/1 線性規劃無法找到解，螞蟻在這樣的情況下仍是維持一定的水準。

六、結論

編製測驗卷問題的研究已有一段時間，目前所提出選題方法均是在數量小的題庫上作挑選，當題庫數量大至數千或者到萬題以上，編製的難度便會大大的提升。螞蟻族群演算法用於解決組合最佳化問題的方法，在 TSP 問題上證明螞蟻演算法所求的解皆好於基因演算法、虛擬退火演算法等等 [16]。因此我們嘗試將螞蟻演算法用於編製測驗卷問題，螞蟻族群演算法所找到的結果於其他的演算

法所找到的結果相差不大，而在時間的花費上少於其他的方法，且當題庫數量變大時，螞蟻族群演算法所找的解仍然能夠維持一定的品質，花費的時間也在一個合理的時間內。由實驗的結果我們相信在編製測驗卷問題上，在眾多選題方法中，螞蟻族群演算法可以是另一個更好的選擇。

七、參考文獻

- [1] 余民寧，「試題反應理論的介紹-測驗理論的發展趨勢(一)」，研習資訊，第 8 卷，第 6 期，民國 81 年，頁 13-18。
- [2] 余民寧，「試題反應理論的介紹-測驗理論的發展趨勢(二)」，研習資訊，第 9 卷，第 1 期，民國 81 年，頁 5-9。
- [3] 余民寧，「試題反應理論的介紹(三)-試題反應模式及其特性」，研習資訊，第 9 卷，第 2 期，民國 81 年，頁 6-10。
- [4] 余民寧，「試題反應理論的介紹(七)-訊息函數」，研習資訊，第 9 卷，第 6 期，民國 81 年，頁 5-9。
- [5] 余民寧，「試題反應理論的介紹(八)-測驗的編制」，研習資訊，第 10 卷，第 1 期，民國 81 年，頁 6-10。
- [6] 李明如，測驗選題演算法之研究—以國中基本學力測驗之國文、數學為例，中原大學資訊工程所碩士論文，民國 91 年。
- [7] 巫眇鶯，運用線性規畫於選題策略之研究，國立台南師範學院資訊教育研究所碩士論文，民國 91 年。
- [8] 孫光天、陳岳宏、賴膺守、謝凱隆、陳新豐，「使用貪婪演算法作為一有效率之選題策略」，全國計算機會議，民國 88 年。
- [9] 孫光天、陳岳宏、賴膺守、謝凱隆、陳新豐，「利用基因演算法於選題策略之研究」，全國計算機會議，民國 88 年。
- [10] 孫光天、程千芬、蔡淑燕，「運用進階基因演算法於選題策略之研究」，全國計算機會議，民國 92 年。
- [11] 黃家輝，利用灰色理論於選題策略之研究，義守大學資訊工程學所碩士論文，民國 91 年。
- [12] 陳健一，應用貪婪式基因演算法於選題策略之研究，暨南大學資訊管理研究所碩士論文，民國 92 年。
- [13] Adema, J. J., Boekkooi-Timminga, E. and van der Linden, W. J., "Achievement test construction using 0-1 linear programming," *European Journal of Operational Research*, vol. 55, pp. 103-111, 1991.
- [14] Chu, P.C. and Beasley, J. E., "A Genetic Algorithm for the Multidimensional Knapsack Problem," *Journal of Heuristics*, vol. 4, no. 1, pp. 63-86, 1998.
- [15] Dorigo, M., Maniezzo, V. and Colomi, A., "The Ant System: Optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics-Part B Cybernetics*, vol. 26, no. 1, pp. 1-13, 1996.
- [16] Dorigo, M. and Gambardella, L., "Ant colony system: A cooperative learning approach to the traveling salesman problem," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 53-66, 1997.
- [17] Dorigo, M., Di Caro, G. and Gambardella, L., "Ant algorithm for discrete optimization," *Artificial Life*, vol. 5, no. 2, pp. 137-172, 1999.
- [18] Fréville, A., "The multidimensional 0-1 knapsack problem: An overview," *European Journal of Operational Research*, vol. 55, pp. 1-21, 2004.
- [19] Hwang, G. J., Hsiao, J. L. and Tseng, J. C. R., "A Computer-Assisted Approach for Diagnosing Student Learning Problems in

- Science Courses,” *Journal of Information Science and Engineering*, vol. 19, no. 2, pp. 229-248, 2003.
- [20] Hambleton, R. K. and Swaminathan, H., Item response theory: Principles and Applications. Kluwer-Nijhoff, 1985.
- [21] Kellerer, H., Pferschy, U. and Pisinger, D., Knapsack problems, Springer-Verlag berlin Heidelberg, 2004.
- [22] Shyu, S. J., Lin, B. M. T. and Hsiao, T. S., “Ant colony optimization for the cell assignment problem in PCS networks,” *Computers and Operations Research*, vol. 33, no. 6, pp. 1713-1740, June 2006.
- [23] Shyu, S. J., Yin, P. Y. and Lin, B. M. T., “Ant-Tree: an ant colony optimization approach to the generalized minimum spanning tree problem,” *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 15, no. 1, pp. 103-112, 2003.