

行政院國家科學委員會補助專題研究計畫成果報告

遺傳演算法應用於資訊專業人員遴選之研究

計畫類別：個別型計畫 整合型計畫

計畫編號：NSC89-2416-H-006-007-SSS

執行期間：88 年 8 月 1 日至 89 年 7 月 31 日

計畫主持人：蔡燿全

共同主持人：

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：國立成功大學企業管理系

中 華 民 國 八十九 年 七 月 三十一 日

行政院國家科學委員會專題研究計劃成果報告

遺傳演算法應用於資訊專業人員遴選之研究

The Application of Genetic Algorithms to Selection of Information Professionals

計劃編號：NSC89-2416-H-006-007-SSS

執行期限：88 年 8 月 1 日至 89 年 7 月 31 日

主持人：蔡燿全>Email:tsaia@mail.ncku.edu.tw 執行機構：國立成功大學企業管理系

計劃參與人員：楊棠堯(yangty@mail.ksut.edu.tw) 執行機構：國立成功大學企管系博士班

計劃參與人員：莊宗南(tnchuang@mail.stut.edu.tw) 執行機構：國立成功大學企管系博士班

一、中文摘要

根據一些研究指出到公元二千年，台灣地區合適之資訊專業人員將呈現嚴重供需失衡的狀態，許多企業也為了遴選合適的資訊專業人員而煩惱。為彌補資訊專業人員短缺的現象，透過正規的教育管道雖可提高人才的供應量，但需花費較長的時間而緩不濟急，因此如何遴選優秀的資訊專業人員乃成為企業界所關注的問題之一。

本研究以遺傳演算法來學習取得人員遴選準則(Criteria)的權重，降低傳統 AHP 方法的主觀成分。另外，傳統人員遴選均只有錄用與不錄用的決策，本研究突破此限制，在每個染色體中，除了各準則的權重外，另加入兩個臨界值(Cut-off value)，使得人員遴選決策更有彈性。最後，本研究結果與 AHP 做一比較，並討論結果隱含的意義。

關鍵詞：遺傳演算法、資訊人員遴選、多準則決策

Abstract

Acute competition in the modern business environment has led more and more companies to employ computer technology in the processing of information. Industrial computerization has led to a growing demand for information professionals in recent years. Therefore, the problem of how to select good information professionals has become an increasingly important issue.

This study addresses the problem through the employment of Genetic Algorithms (GA). These are used to generate the weightings of various selection criteria

for information professionals, and thereby reduce the subjectivity of the AHP method. In addition to the traditional "yes or no" decisions for the selection of information professionals, our approach uses two cut-off values for each chromosome, which represent the weighting values, in order to make the selection decisions more flexible. The results of using our approach are compared with the AHP method, and some practical implications are discussed.

Keywords: Genetic Algorithms, Selection of Information Professionals, Multicriteria Decision

二、緣由與目的

Gallagher(1988)將管理資訊系統的發展劃分為三個時期，第一是資料處理期：主要目的是作業支援，強調效率提昇；第二是管理支援期：主要目的是輔助管理，強調效能提昇；第三是競爭態勢期：主要目的是加強競爭優勢，強調市場佔有率與獲利率。由此可見，過去資訊系統(IS)只是被公司用來作為資料處理與決策支援的輔助工具，但隨著競爭環境的激烈與資訊科技的快速發展，資訊系統所扮演的角色已不再只是用來支援組織企業活動的工具，更是企業組織用來提昇競爭力的一項利器(Bakos & Treacy, 1986; Porter & Miller, 1985; Wiseman, 1985)。因應日趨激烈的競爭環境，越來越多的行業以電腦科技來處理資訊，而資訊電腦化之後，資訊專業人員的需求亦逐年增加，並且已經到了供不應求的情況。根據一些研究指出到公元二千年，台灣地區合適之資訊專業人員將呈現嚴重供需失衡的狀態，許多企業也為了

遴選合適的資訊專業人員而煩惱。為彌補資訊專業人員短缺的現象，透過正規的教育管道雖可提高人才的供應量，但需花費較長的時間而緩不濟急，因此如何遴選優秀的資訊專業人員乃成為企業界所關注的問題之一。

企業界不管是現在或未來，都會面對資訊人力資源短缺，且適當就業人才難覓的窘境。加上資訊科技的快速變遷，面對市場競爭日趨激烈的企業，實有必要瞭解公司本身的資訊需求，積極培育與充實資訊專業人員的技能與知識；另一方面，企業應建立一套遴選制度，審慎評選新進人員所具備專業技能與知識。而由於資訊專業人員工作性質的不同，其遴選的條件也和一般的企業員工有所差異。傳統衡量資訊人員的方式，除了考慮其專業能力外，更應該從各個角度來評估其潛力。目前用來判斷資訊專業人員能力的方式，仍然是以傳統的二值邏輯觀念來陳述，所以在能力好壞的比較上常常會有模糊、不確定的現象產生，因此希望能以更客觀有效的方式來遴選優秀的資訊專業人員。

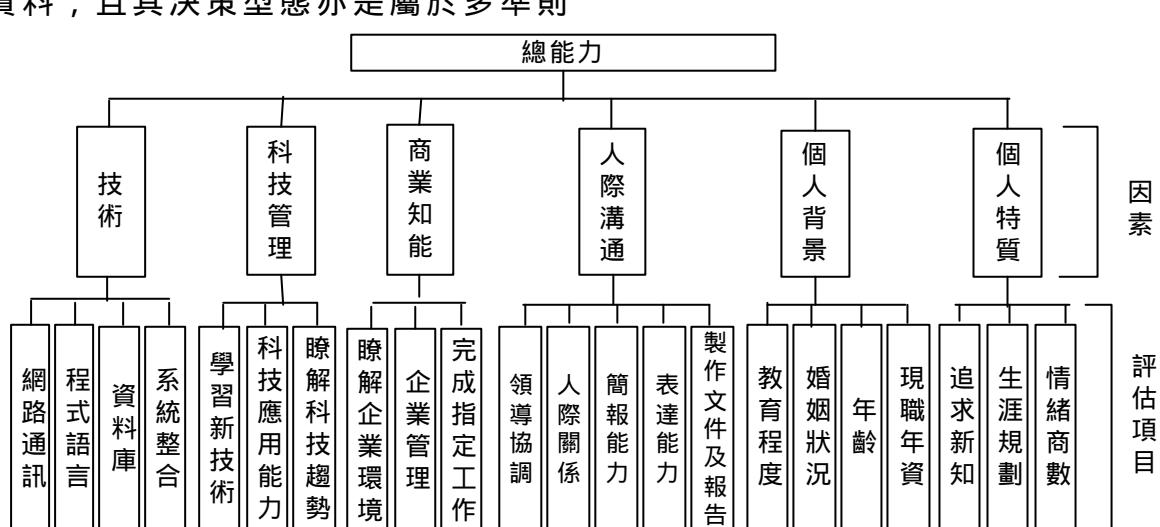
資訊專業人員的遴選並不像其他具有數量化的決策模式可以直接套用，因決策問題本身的變數都是屬於定性的非數量化資料，且其決策型態亦是屬於多準則

(Multiple criteria)決策，所以傳統的解決方式大部分都是使用層級分析法（鄧振源、曾國雄，1989；Zahedi, 1996）。然而由於層級分析法在成對比較時，受訪者可能會有相當大的主觀偏差，且求取權重亦是採用簡單的加權平均方式，如此導致了學者對層級分析法的主觀成分產生了質疑。

為了能解決上述問題，本研究以遺傳演算法來學習資訊專業人員遴選準則的權重，並加入兩個臨界值(Cut-off value)的學習，使得遴選決策能有錄用、備取、不錄用的彈性，更加強實際的應用價值。在遺傳演算法的交配(Crossover)運算中，本研究採用所謂連續一致性交配演算法(Continuous uniform crossover)，來確保下一代產生的染色體均是合法值。

三、研究架構

根據文獻整理出的各項影響資訊專業人員遴選的因素及評估項目後（林東清；湯宗泰，1995；吳奇才等，1998；陳建宏，1996；張緯良等，1994；Denis et. al, 1995），本研究採取的22個評估項目架構如圖一所示，其權重代號依序為網路通訊(w_1)、程式語言(w_2)、資料庫(w_3)、生涯規劃(w_{21})、情緒商數(w_{22})。



圖一、資訊專業人員遴選之準則架構

本研究在資訊業、服務業、製造業等不同產業共發出500份問卷，對象均為公公司的資訊部門主管，回收93份，剔除無效問卷8份，有效樣本計85份，亦即本研究GA模式的母體有85筆染色體。當問卷回收後，先利用AHP的演算法求得每份問卷中個別評估項目的權重值，這些權重值即

為每個染色體的基因值(Allele)。

遺傳演算法(GA)是一種新提出的Stochastic搜尋演算法，意即凡是使用機率來輔助問題的最佳解搜尋者，均屬此種演算法(Grant, 1995)。遺傳演算法是John Holland於1970年提出，如同其名，此演算法的推論過程就好像生物學的基因繁

殖。遺傳演算法將資料編碼成一連串的資訊，如同 DNA 的染色體，染色體間透過複製（母代染色體不經過任何處理，整個染色體保留到下一子代）交配（從母代選取出來的染色體，由這兩個染色體中的某部分基因互相對調，形成新的兩個染色體，而產生新子代）突變（母代染色體中，部分基因發生改變）等行為，根據最適者生存、不適者淘汰的規則，最後獲得最佳的染色體，而此染色體即為問題的最適解（Goldberg, 1989）。

本研究染色體的編碼是由 22 個評估項目的權重值，再加上 2 個臨界值，共 24 個基因所組成。每個基因均採實數值編碼，且其值均介於 0 與 1 之間。假如遴選準則有 k 個，則染色體(c)的編碼如下式所定義：

$$c = \langle w_1, w_2, \dots, w_k, x_1, x_2 \rangle$$

其中 w_j 代表第 j 個準則的權重，且

$$\sum_{j=1}^k w_j = 1；而 x_1 與 x_2 分別為上臨界值與下$$

臨界值，且 $x_1 > x_2$ 。

而最後用來判定遴選決策則是使用權重總和(Weighted sum)，假設有一資訊專業人員(i)，其染色體(c)，則權重總和 $ws(c,i)$ 定義如下：

$$ws(c,i) = \sum_{j=1}^k w_j \frac{i_j - \min_j}{\max_j - \min_j}$$

其中 i_j 是此人員 i 的第 j 個準則的權重值， \max_j 及 \min_j 則代表所有人員的第 j 個準則的最大及最小權重值。根據上述的權重總和，當 $ws(c,i) \geq x_1$ ，則遴選決策為『錄用』；當 $x_2 \leq ws(c,i) < x_1$ ，則遴選決策為『備取』；當 $ws(c,i) \leq x_2$ ，則遴選決策為『不錄用』。

染色體的適合函數(Fitness function)是用來反應正確遴選的能力，因此，任何遴選錯誤的染色體均會被加上修正值 P_i ，其值如下：

$$P_i = \begin{cases} 1, & selection(i,c) = selection(i) \\ 0.4, & |selection(i,c) - selection(i)| = 1 \\ 0, & otherwise \end{cases}$$

其中 $selection(i)$ 是由公司主管所做的第 i 個人員遴選決策，而 $selection(i,c)$ 則是 GA 模式所做的遴選決策。所以染色體的適合函數定義如下：

$$fitness(c) = \frac{1}{t} \sum_{i=1}^t P_i$$

其中 t 代表染色體母體的個數。

當系統所做的遴選決策與公司主管所做的遴選決策相同時，則 $fitness(c)$ 的值為 1，當所有染色體均與公司主管所做的遴選決策不相同時，則 $fitness(c)$ 的值為 0，所以 $fitness(c)$ 的值是介於 0 與 1 之間。

在遺傳演算法中，最重要的是交配(Crossover)運算，通常是以一固定機率(P_c)從母體中隨機選出一對染色體來做交配運算，模擬過程需考量 P_c 的大小， P_c 過高會造成父代較佳特性的流失， P_c 太低則會失去進化的目的。最基本的交配運算是單點交配(Single-point crossover)，其運算方式如表一所示：

表一、單點交配

父代	子代
父 1 : 10011011	子 1 : 10011101
父 2 : 01001101	子 2 : 01001011

但傳統的單點交配運算在本研究可能會造成錯誤的編碼，例如，假設有下列兩個染色體：

$$x = \langle 0.1, 0.4, 0.3, 0.2 \rangle$$

$$y = \langle 0.1, 0.3, 0.1, 0.5 \rangle$$

則單點交配運算會產生下列的子代：

$$x' = \langle 0.1, 0.4, 0.3, 0.5 \rangle$$

$$y' = \langle 0.1, 0.3, 0.1, 0.2 \rangle$$

因 $x' = 1.3 > 1$ 且 $y' = 0.7 < 1$ ， x' 與 y' 均沒有符合前述條件 $w_j = 1$ ，均是錯誤的編碼。

為了解決上述問題，本研究採用所謂連續一致性交配(Continuous uniform crossover)運算，其能產生合法的子代編碼，其定義如下：

假設有兩個染色體 $x = \langle x_1, x_2, \dots, x_n \rangle$ 及 $y = \langle y_1, y_2, \dots, y_n \rangle$ ，其子代定義為 $x' = \langle x'_1, x'_2, \dots, x'_n \rangle$ 及 $y' = \langle y'_1, y'_2, \dots, y'_n \rangle$ ，則

$$x'_i = sx_i + (1-s)y_i$$

$$y'_i = (1-s)x_i + sy_i$$

其中 s 是一個常數，本研究是以 $[-0.5, 0.5]$ 來隨機選擇 s 值。此運算保證子代的權重基因值總和為 1，證明如下：

假設 $\sum_{i=1}^m x_i = 1$ ，則

$$\sum_{i=1}^m x'_i = s \sum_{i=1}^m x_i + (1-s) \sum_{i=1}^m y_i = s + (1-s) = 1$$

而 $\sum_{i=1}^m y'_i = 1$ 亦成立

而且連續一致性交配亦保留基因值之間 greater-than 的關係，因每個染色體均有兩個臨界值 x_1 及 x_2 ，假如 $x_1 > x_2$ ，則 $x'_1 > x'_2$

恆成立。

其它如遺傳代數的設定，本研究採用設定固定值方式，但為求準確率較高，故設定為 1000。而突變(Mutation)運算的值設定為 0.001。

綜上所述，表二列出本研究 GA 模式的各項設定值：

表二、遺傳演算法參數

參數名稱	設定值
染色體長度	24
母體大小	85
遺傳代數	1000
一般複製比例	0.05
交配機率	0.7
突變機率	0.001

四、結果與討論

本研究使用的 GA 軟體是 Evolver for EXCEL 97，各參數的設定值如上述表二，染色體母體大小為 85，染色體長度為 24，表現型式為 $\langle w_1, w_2, \dots, w_{22}, x_1, x_2 \rangle$ ，交配機率 (P_c) 為 0.7，突變機率 (P_m) 為 0.001，經過演算過程，在第 132 代時，適合函數值收斂在 1 的地方，其每一代的適合函數值如圖二所示。



圖二、適合函數值

最後 GA 模式學習得到的資訊專業人員遴選準則的最佳權重值 (w_1, w_2, \dots, w_{22})，再加上兩個臨界權重值 (x_1, x_2)，便可用來模擬新進資訊專業人員的遴選決策。當得到另一組樣本權重值，透過下式的計算，

$$ws(c, i) = \sum_{j=1}^k w_j \frac{i_j - \min_j}{\max_j - \min_j}$$

可以得到每個染色體的權重總值 $ws(c, i)$ ，再與 x_1 及 x_2 來比較，便可得到每位樣本的遴選結果，其決策規則為：

```

IF ws(c,i)>= x1 THAN
    遴選結果=『錄用』
ELSEIF x2<=ws(c,i)< x1 THAN
    遴選結果=『備取』
ELSE
    遴選結果=『不錄用』
END IF

```

表三是 GA 演算法學習後的各準則的最佳權重值 (w_1, w_2, \dots, w_{22}) 及上臨界權重值 (x_1) 與下臨界權重值 (x_2)，其 $\sum_{j=1}^{22} w_j = 1$ 。同樣的樣本，以傳統的 AHP 方法來求得各準則的權重值，兩者比較於表三。

由表三可以看出，網路通訊 (w_1) 的權重值是 0.1889，乃所有評估項目準則中最高的，由於本研究的對象是各產業的資訊專業人員的遴選，而現今網際網路 (Internet) 蓬勃發展之際，各產業亦紛紛投入 B2B 或 B2C 的電子商務競爭，在需求面來看，本研究的結果無異與現今產業需求不謀而合。然而，以 AHP 計算得到的結果，卻是瞭解科技趨勢 (w_7) 的權重值最高，但 GA 模式的瞭解科技趨勢 (w_7) 準則亦高達 0.1081，顯示在現今資訊科技的快速發展下，日新月異的資訊科技新產品，在公司的資訊部門主管眼中，是企業取得競爭優勢的重要因素。

表三、GA 與 AHP 之權重值比較

GA		AHP	
w ₁ =0.1889	w ₂ =0.0859	w ₁ =0.0942	w ₂ =0.0879
w ₃ =0.0936	w ₄ =0.0455	w ₃ =0.0561	w ₄ =0.0714
w ₅ =0.0405	w ₆ =0.0463	w ₅ =0.0571	w ₆ =0.0734

$w_7=0.1081$	$w_8=0.0286$	w₇=0.1156	$w_8=0.0104$
$w_9=0.0108$	$w_{10}=0.0526$	$w_9=0.0544$	$w_{10}=0.0516$
$w_{11}=0.0259$	$w_{12}=0.0333$	$w_{11}=0.0274$	$w_{12}=0.0564$
$w_{13}=0.0097$	$w_{14}=0.0111$	$w_{13}=0.0031$	$w_{14}=0.0561$
$w_{15}=0.0019$	$w_{16}=0.0025$	$w_{15}=0.0123$	$w_{16}=0.0108$
$w_{17}=0.0005$	$w_{18}=0.0040$	$w_{17}=0.0085$	$w_{18}=0.0474$
$w_{19}=0.0009$	$w_{20}=0.0958$	$w_{19}=0.0005$	$w_{20}=0.0245$
$w_{21}=0.0681$	$w_{22}=0.0455$	$w_{21}=0.0358$	$w_{22}=0.0451$
$x_1=0.476$	$x_2=0.325$	$x_1=0.675$	$x_2=0.335$

另外，在個人背景因素（包括教育程度 w_{16} 、婚姻狀況 w_{17} 、年齡 w_{18} 、現職年資 w_{19} ）中，GA 演算法學習後的權重值是最低的，但 AHP 方法所得到的權重值則是平均大小，此乃可能樣本中有些非資訊業的產業，其主觀成分較重，而影響 AHP 求得的結果，這也可以說明本研究的遺傳演算法人員遴選模式較能減少主觀認定造成的影響。

本研究與文獻上其它方法的人員遴選模式最大的不同是，採用臨界值的比較來進行遴選決策。傳統使用 AHP 或模糊綜合評判等方法，僅能將所有人員依權重成績 (Weighted score)——排序，錄取或不錄取的界線亦採主觀認定；但本研究的遺傳演算法人員遴選模式則可以依個別人員的特質來進行遴選，較能實際應用於企業聘僱人才的決策輔助。然而，本研究美中不足的是，因採用郵寄問卷方式來蒐集樣本，無法控制受測者的身份及填答時的情況，研究結果可能誤差較大，未來應採用實地訪問方式來取得更確切的樣本資料，以提高本研究的應用價值。

五、參考文獻

- [1] 鄧振源、曾國雄，「層級分析法(AHP)的內涵特性與應用（上）」，中國統計學報，第 27 卷，第 6 期，第 5-22 頁，民國 78 年 6 月。
- [2] 鄧振源、曾國雄，「層級分析法(AHP)的內涵特性與應用（下）」，中國統計學報，第 27 卷，第 7 期，第 1-20 頁，民國 78 年 7 月。
- [3] 林東清，「資管人員的技能需求—由組織管理、系統開發與電腦科技三方面來比較分析」，資訊管理，第一卷，第一期，第 60-77 頁。
- [4] 湯宗泰，「資訊管理專業技能需求面與供給面之比較研究」，第七屆中華民國管理教育研討會，第 191-198 頁，民國 84 年。
- [5] 吳奇才、黃美文、蔡瑞明，「資訊系統新進人員技能之需求分析」，第九屆國際資訊管理學術研討會，民國 87 年。
- [6] 陳建宏，「資管人員情緒商數的研究」，大葉大學碩士論文，民國 85 年。
- [7] 張緯良、王贊旭，「資訊專業人員個人背景特徵與工作績效之研究」，人力資源學報，第四期，第 83-94 頁，民國 83 年。
- [8] Gallagher, J. P., "Knowledge systems for business", Prentice-Hall, Englewood Cliffs, 1988, pp. 7-8.
- [9] Bakos, J. Y. and M.E. Treacy, "Information technology and corporate strategy: A research perspective," MIS Quarterly, vol. 10(2), 1986, pp. 107-119.
- [10] Porter, M. E. and V.E. Miller, "How information gives you competitive advantage," Harv. Bus. Rev., vol. 63(4), 1985, pp. 149-160.
- [11] Wiseman, C., "Strategy and Computers: Information Systems as a Competitive Weapon", Dow Jones-Irwin, Homewood, 1985.
- [12] Zahedi, F., "The Analytical Hierarchy Process-A Survey of the Method and its Application", Interfaces, Vol. 16, No. 4, 1996, pp. 96-108.
- [13] Denis, M. S. and Lee, "Critical Skills and Knowledge Requirements of IS Professionals: A Joint Academic/Industry Investigation", MIS Quarterly, 1995, pp.313-340.
- [14] Grant, K., "An Introduction to Genetic Algorithms", C/C++ Users Journal, 1995, pp.45-58.
- [15] Goldberg, D. E., "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, MA, 1989.