

行政院國家科學委員會專題研究計畫 成果報告

網路論戰與敵意之研究：認知歷程、情境線索與自動化文章 敵意分級系統(2/2)

計畫類別：個別型計畫

計畫編號：NSC93-2520-S-009-001-

執行期間：93年08月01日至94年12月31日

執行單位：國立交通大學教育研究所

計畫主持人：林珊如

共同主持人：劉旨峰

計畫參與人員：黃宏宇、劉怡秀、林淑卿、蔡婷、陳滢方、林建妤、林志鴻

報告類型：完整報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 94 年 10 月 17 日

行政院國家科學委員會補助專題研究計畫 ^{結案} 報告

計畫名稱

網路論戰與敵意之研究：認知歷程、情境線索與
自動化文章敵意分級系統(1/2)(2/2)

計畫類別：個別型計畫

(1/2)計畫編號：NSC 92-2520-S-009-005-

(1/2)執行期間：2003 年 8 月 1 日至 2004 年 7 月 31 日

(2/2)計畫編號：NSC 93-2520-S-009-001-

(2/2)執行期間：2004 年 8 月 1 日至 2005 年 12 月 31 日

計畫主持人：林珊如

共同主持人：劉旨峰

計畫參與人員：黃宏宇、劉怡秀、林淑卿、蔡婷、陳滢方、林建妤、
林志鴻

成果報告類型(依經費核定清單規定繳交)：結案報告

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢

執行單位：交通大學教育研究所

中 華 民 國 94 年 10 月 10 日

計畫名稱：網路論戰與敵意之研究：認知歷程、情境線索與自動

化文章敵意分級系統(1/2)(2/2)

計畫主持人：林珊如

(1/2)計畫編號：NSC 92-2520-S-009-005-

(1/2)執行期間：2003年8月1日至2004年7月31日

(2/2)計畫編號：NSC 93-2520-S-009-001-

(2/2)執行期間：2004年8月1日至2005年12月31日

執行機關：國立交通大學教育研究所

為何要研究網路的文字攻擊、檢驗引發攻擊的歷程、並發展緩和衝突的教學與訓練方案呢？首先從1980年以後已有文獻記錄，不論是在公司學校等工作情境、或以休閒聊天為目的所進行的各式網路中介溝通，火爆論戰(flaming)是經常見到的(1980-1989年：如 Siegal, Dubrovsky, & McGuire, 1983; Baron, 1984; Kiesler, S., Siegel, & McGuire, 1984; Hiltz, Turoff, & Johnson, 1989；1990-to date: 如 Lea, O'Shea, Fung, & Spears, 1992; Thompson, & Ahn, 1992; Tompson & Fulger, 1996; Reinig, & Brigg, 1997; Mabry, 1998; Wallace, 1999; Milam, & Waller, 2000)。以台灣使用者為樣本的研究則直到1999年之後才出現：如張千駿(1999)、張紳震(2000)、甘百瑩(2002)、倪珮晶(2003)、黃宏宇(2003)、劉怡秀(2004)、及林珊如與黃宏宇(2004)。其中大部分都以大眾傳播的角度進行論戰文字的內容分析，只有少部份研究觸及論戰中個人所經歷的憤怒或敵意內在心理歷程(黃宏宇, 2003; 劉怡秀, 2004; 及林珊如與黃宏宇, 2004; Ybarra & Mitchell, 2004)，也只有少數研究，如倪珮晶(2003)分析文字衝突的策略。

相關的研究有時因為研究者的角度不同，而採用不同的變項名稱，例如近年來美國的老師、家長、與諮商人員由於關心網路的安全性，發現兒童及青少年期經常發生的欺侮(bullying)事件，原本只出現在遊戲區或學校中成人監督不到的場所，包含身體推擠、口語的揶揄或恐嚇，如今逐漸由現實世界帶入網路討論區，稱為網路欺侮(online bullying or cyberbullying, www.bullyonline.org, 2004; www.bewebaware.org, 2004)，在網路上無法進行身體攻擊，只能利用語言來進行情節輕重不同的攻擊，包含：排擠、鄙視、扭曲語意、製造人身攻擊的謠言、與恐嚇，甚至在當事人被迫離開某一討論區之後，還有跨網站或跨討論區的網路跟蹤，所利用的是欺侮者與被欺侮者間權力地位的不平衡，同時欺侮的對象已經不限於青少年，越來越多成年人發現自己因為一段不經意輸入的語句，而成為被欺侮者。而 Suler(2004)的網路心理學網站，也因為虛擬世界的攻擊極為盛行，特闢一節描述網路的衝突(conflict online)。本人則從90年起以心理學的敵意(hostility)或文字攻擊(verbal aggression)作為概念定義，受國科會科科教處的經費支助，研究網路上

的攻擊。本研究的網路文字攻擊不包含靜態張貼帶有恨意或鄙視(種族、性別、社經地位等)的文字，亦即不只是等待觀眾來閱讀的敵意文字，而特指網路討論區的人群間所發生的文字衝突，重點是要有「主動的參與者」，也特別著重主動參與者的心理歷程與影響的行為因素。

到底網路上的文字攻擊(在台灣)有多盛行呢？

本人曾經試圖對我國各大學 BBS 站的各種討論區蒐集資料，以便對俱攻擊性的佈告(posts)篇數進行描述性統計，但是蒐集了三個月的資料後，發現難以克竟全功，因為許多大型 BBS 站的佈告篇數以每周約 1,000 到 10,000 篇的速度增加，達到電腦儲存的空間上限時，就由系統自動依張貼的時間予以刪除，所以 BBS 版的佈告之母群是每周(天)變動的。論戰的佈告數隨校園活動的對立性、運動季節、政治議題、競選時間、連續劇或電視節目的流行、明星的活動、或寒暑假學生休息時間而有非常劇烈的起伏。各大學的 BBS 電子佈告欄大都具有全國轉信的功能(視站長的設定，少部分討論區不轉信)，打破了現實世界的地理區隔，使來自不同學校的使用者能對相似的議題發表意見、進行討論，同時每個 BBS 站有不同的版規，有的站或版會刪除文字過度火爆或咒罵性的佈告，有的則以言論自由為尚，因此難以取決應該如何取樣，而如果進行單一時間點的抽樣，本人對該描述性統計的可信度與可推論性都不俱信心。

但是東華大學高台倩教授所指導的倪珮晶同學(2003)提供了一個概略的推估，以攻擊文字盛行的六個討論區：棒球、電視、政治、宗教與心靈、交通大學校園、清華大學校園為範圍，在 92 年 2 月到 4 月間的兩個半月間，每五天抽樣一次，進行 14 次取樣，取樣範圍必須是參與者為三人以上、同一主題的討論佈告在五天內達到五篇以上的討論串。由兩位編碼員以對論戰的激烈程度給予 1-7 級的評分(友善、中立、分歧、爭論、緊張、敵對、攻擊)，唯有第三級(含)以上的佈告才計為論戰佈告。結果發現有 1469 篇被列為論戰佈告，分別屬於 50 個主題。也就是在 50 個論戰主題，佈告是以討論串的方式出現，少則為 10 篇的討論串，多的甚至有 100 篇以上結成討論串—討論同一主題，而她蒐集的論戰討論串平均有 29.38 篇佈告(中位數 18.5 篇)，標準差是 29.42。遺憾的是由於佈告母群的變動性，她也無法提供論戰佈告篇數相對於總佈告篇數的比例值。就論戰時間而言，有些討論串在一小時內就結束了，有的長達 274 小時(約 12 天)，但大多數論戰討論串(72%)在三天內結束。至於討論串的論戰激烈程度，有 36%屬於第三級分歧、有 44%為第四級爭論，而比較激烈的第五級緊張(18%)與第六級敵對(2%)的比例則急速下降。

其實只要是論戰的文字一經張貼到 BBS 站上，觀看到該文字的人數根本無法計算，而對觀看者(不見得會回應，即所謂的 lurker)的心理影響力更是難以估計，有的人覺得事不關己，有的人則深受傷害，也有許多人覺得這就是網路的次文化。因此文字的內容分析不足以反應觀察者或參與者(攻擊者 aggressors 與受害者 victimizers)的心理感受。Ybarra & Mitchell (2004)以美國全國的 10-17 歲青少年為母群，以電話訪問隨機抽樣的 Internet 固定使用者，得到 1501 份資料。樣本中有 19%自陳參與過網路的文字攻擊，其

中有 4% 是被攻擊者、有 12% 是攻擊者、也有 3% 交互擔任過被攻擊與攻擊者，該研究並未計算「不參與、而只是觀看」文字攻擊的人數。其次是，這些人有類似欺侮者 (the bully) 及被欺侮者 (the victim) 的社會心理現象，例如欺侮人的與被欺侮的之間，在所處社群中存有權力與地位的不平衡，會欺侮別人的掌控住比較多權力，其特性是比較有自信、有攻擊性、對受欺侮的人沒有同情心，而被欺侮的通常比較退縮、害羞、沒有朋友的支持、無法有效因應別人的攻擊，被人欺侮的可能出現長期性的負面影響，例如喪失自尊心、在人前容易害怕、恐懼、憤怒、不信任人，對人際交往會越來越沒自信。

以上的文獻記錄與本人的研究經驗顯示，網路討論區中的文字攻擊數量極多，美國青少年主動參與的比例約為 20%，我國的主動參與人數則尚未能估計出來，網路世界反映的不過就是一般人性與人類行為，而以人類日常生活的攻擊行為極為盛行來看，網路文字攻擊的現象不會比現實世界少，因此值得發展心理學者、教育學者、資訊科學研究者共同觀察或監督。

民 91 年研究成果：為網路敵意與攻擊的研究奠基

本人在 90-91 年以心理學的敵意 (hostility) 為概念定義，進行網路社會環境的去抑制研究 (social disinhibition, 國科會類教處支助，計畫編號 NSC 90-2511-S009-006)。此為第一年的初探性研究，樣本限定為國立交通大學學生，透過網路收集問卷資料，有效樣本 464 份，其中男性 333 份，女性 131 份。文章已刊登於 TSSCI 期刊：教育與心理研究，2004, 27(2), 頁 325-352。

在研究工具方面，編製了網路敵意量表 (Internet Hostility Scale, IHS)，用以分析面對論戰的內在個人感受，以探索性因素分析萃取出六個分量表，因素含：

1. 網路敵意的行為面，有四個因素：「其他網路攻擊手段」、「直接口語攻擊」、「運用團體力量攻擊」、「忽視權利」，前二者表現直接明顯的攻擊，後二者為迂迴的攻擊。
2. 網路敵意的情意層面為「表現忿怒」因素。
3. 網路敵意的認知層面即「網路人際懷疑」因素。

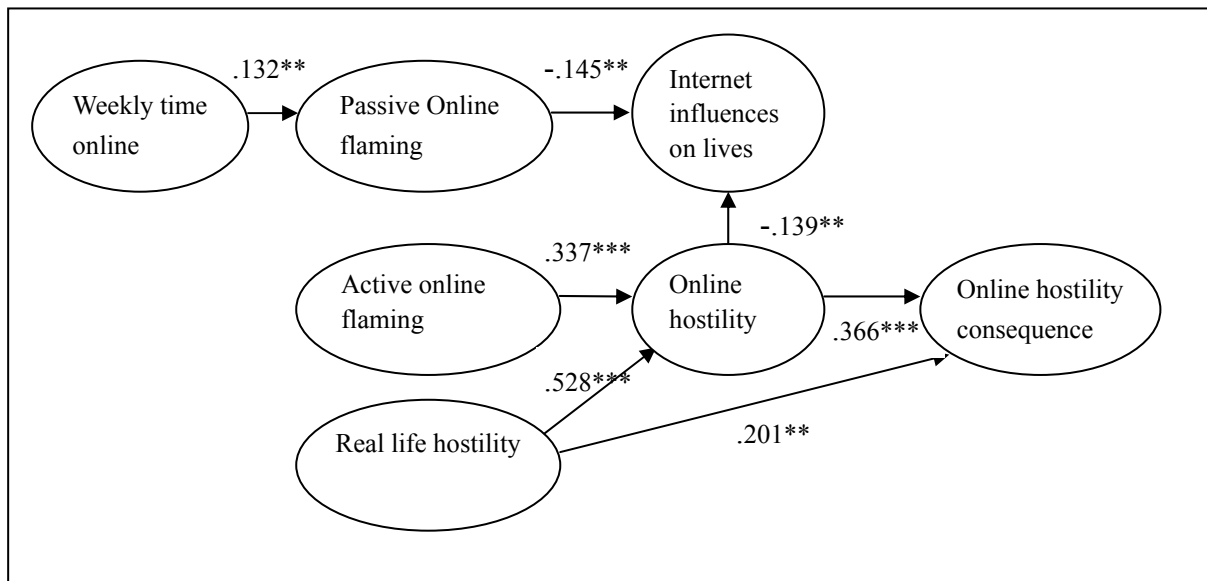
本研究結果肯定以往大眾對網路論戰的直觀經驗：BBS 經常有一群人在爭吵與衝突，常見到衝突爭吵的版面多為公共論壇型的版面，如校園版、抱怨版、研究生版、匿名版。受試者多半是在 BBS 上觀看到他人衝突的事件，自己參與成為主動攻擊者或遭受攻擊的狀況較少。此結果與國外對網路論戰的研究相符：「閱讀網路張貼文字或線上文字交談時，比面對面溝通來得容易誤解對方具有敵意，而爆發怒氣或攻擊」(Hiltz, Turoff, & Johnson, 1989; Mabry, 1998; Siegal, Dubrovsky, & McGuire, 1983)。同時，網路衝突經驗頻率愈高者，網路敵意較高，男性的網路敵意比女性來得高。其次，網路沉迷量表篩選出來的高分者，其網路使用時數遠高於一般使用者，男性的網路沉迷程度及人數比例高於女性，網路沉迷程度越嚴重的，越是認為網際網路對生活有負面影響，而網路

敵意的得分也愈高。

如果把觀看別人論戰之被動經驗與本人參與論戰(active online flaming)的主動經驗區分開來，則本研究對網路去社會抑制現象之發現見圖一的路徑分析：

1. 長時間使用網路會提高個人觀看到網路論戰的機會(Beta = 0.132)，觀看到網路論戰的經驗使個人評估的網路對生活影響力持負面看法(Beta = -0.145)，但是僅只是觀看論戰的經驗並不能預測網路敵意。
2. 比較特別的是網路使用時間與主動參與論戰之間並無關係，但是主動參與論戰能有效預測個人的網路敵意程度(Beta = 0.337)，網路敵意同時受到過去自己主動參與論戰的經驗及個人在現實世界持有的敵意(Beta = 0.528)之影響。
3. 網路敵意會帶來後續影響，包含對自我概念受損及負面情緒蔓延(Beta = 0.366)。

Joinson(1998)主張論戰是一種網路上的去社會抑制行為(disinhibition，或譯為社會解禁現象)，也有許多學者主張網路是一個比現實世界更具去社會抑制的情境，但是有這樣的情境不必然引發敵意，個人觀看網路上的論戰文字時，不見得會動怒，只有當個人評估網路人群為不友善、產生認知扭曲，或文字激發個人感受到不舒服或受攻擊，則主觀的敵意才會勃然而發，此種網路敵意與現實世界的敵意或許有相似的效果，例如敵意高的人比較容易罹患心血管疾病，死亡率與癌症的罹患率增高 (Barefoot, Dahlstrom & Williams, 1983; Williams, 1991; Smith, 1992)。



圖一：主動參與及被動觀看網路論戰導致網路敵意與後果的路徑分析圖

民 92 年研究成果：找出網路攻擊的人群及影響攻擊的認知與環境變項

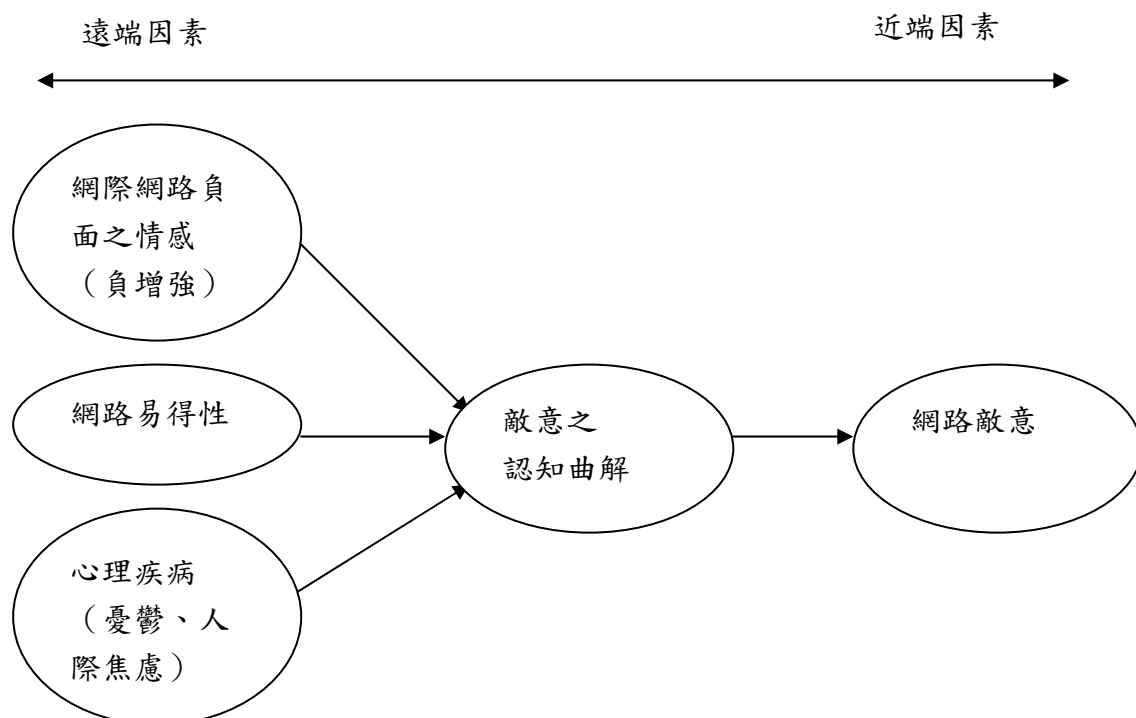
91-92 年的研究主要目的想把網路敵意的心理建構作進一步的澄清，也要檢驗環

境、個人心理特性與認知的中介歷程。因此以兩所公立大學與一所私立大學為樣本，以網路問卷進行調查，樣本共 987 人，包含台灣大學學生 215 份 (21.8%)，交通大學 526 份 (53.3%)，淡江大學 246 份 (24.9%)。男性 625 名 (63.3%)，女性 362 (36.7%)。大學部學生 749 名 (75.9%)，碩士班學生 214 名 (21.7%)，博士班學生 24 名 (2.4%)。

研究問題如下：

1. 高低網路敵意群體在認知行為模式中的遠端與近端變項有差異嗎？其遠端變項有三個：包含環境帶來的「網路負面情感」、「網路易得性」、及個人內在的不適應特性—「心理病理的徵狀（憂鬱、人際焦慮）」，近端變項為個人的認知歷程「認知曲解」，依變項為網路敵意。詳見圖二。
2. 以網路敵意為效標變項是否能區分出數種不同意義的群體？
3. 網路敵意是否具有一認知行為模式？

本研究沿用 90 年的網路敵意操作型定義，並概念定義網路敵意的認知曲解 (cognitive distortion) 為一種不真實、錯誤的態度、武斷或極端化的想法，極端個人中心、僵硬且強制性、將不良認知加以合理化，以非邏輯的方式，產生自我貶抑的評價，加上低度的挫折容忍力，影響個體對於是非的判斷，而表現於外在的行為過程中 (黃宏宇，民 91)。修訂連秀鸞 (民 88) 的「認知扭曲量表」，使之符合網路情境，形成新的網路敵意認知扭曲問卷，共分成四個分量表：負向思考、錯誤標籤、自我中心、責備他人。



圖二 網路敵意之認知行為模式

為回答研究問題 1，將網路敵意問卷全部樣本，依據得分高低區分為極端組，得分前 33% 定義為高網路敵意群體 (n=330)，得分在後 33% 定義為低網路敵意群體 (n=342)，並比較其在敵意認知曲解、網路論戰經驗頻率、網路使用時間、人際焦慮與憂鬱等變項上之差異。研究結果顯示：

1. 高低網路敵意群體在「敵意認知曲解」的多變量變異數分析發現有顯著差異。在各認知扭曲的各向度也發現，高網路敵意群體相對於低網路敵意群體，在負向思考 ($t=-13.00, p<.001$)、錯誤標籤 ($t=-14.12, p<.001$)、自我中心 ($t=-16.54, p<.001$)、責備他人 ($t=-12.83, p<.001$) 與敵意認知曲解總分 ($t=-19.04, p<.001$) 之扭曲程度來得嚴重。
2. 高網路敵意群體相對於低網路敵意群體，無論在「目擊網路論戰經驗頻率」($t=-5.09, p<.001$)、「經歷網路論戰經驗頻率」($t=-13.51, p<.001$) 與「網路論戰經驗頻率」總分 ($t=-13.29, p<.001$)，有比較多的網路論戰經驗。
3. 高網路敵意的群體相對於低網路敵意群體，在心理病理的徵狀：「憂鬱」($t=-3.96, p<.001$) 與「人際焦慮」($t=-5.44, p<.001$) 得分較高。但是在「網路使用時間」方面，不論是每天網路使用時間 ($t=-.74, p>.05$)，或每週網路使用時間 ($t=-1.04, p>.05$) 上，皆無顯著差異存在。因此網路敵意與網路使用時間的關係並不明顯，屬於「量」上網時間，與屬於「質」的網路論戰經驗與憂鬱、人際焦慮等心理歷程相較，質會比量更能影響網路敵意。

為回答研究問題 2，以二階段法進行集群分析。首先隨機從全部樣本中抽取 188 名樣本 (約為五分之一)，以華德法 (Ward's method) 進行階層分類，並由樹狀圖等相關資料判別適當的群體數，結果 188 名樣本可區分成三群體。各群體在網路敵意各向度的差異性檢定上皆達顯著水準，顯示該分類具有效度存在。因此以全體樣本進行三個群體的集群分析，經事後比較將這三群加以命名，如下：

1. 第一群在「忽視權利」與「運用團體力量攻擊」上是屬於高分群，研究者認為此類型的網路使用者較傾向使用迂迴、非直接的方式來表達敵意，因此將命名為「迂迴表達網路敵意群體」，共 468 名(47.4%)。
2. 第二群在「網路人際懷疑」、「表現忿怒」、「直接口語攻擊」、「運用團體力量攻擊」與「其他網路攻擊手段」皆較其他群體來得高，屬於直接、衝動表達敵意傾向，因此命名為「直接表達網路敵意群體」，共 255 名(25.8%)。
3. 第三群體在網路敵意各向度上皆屬於低分群，網路敵意表現不明顯，因此命名為「低網路敵意群體」，包含 264 名(26.7%)。

以區別分析 (discriminant analysis) 進一步檢驗，產生兩組區別函數，都達到顯著水準 (典型相關 η 分別為 .599 及 .322, Wilk's Λ 值分別為 .575 與 .896, $p<.001$)。對區別函數一的影響力較大的為「主動參與網路論戰經驗」、「自我中心」、「負向思考」、「病態網路使用」，以函數一可以把第一群「迂迴表達網路敵意群體」與第二群「直接

表達網路敵意群體」區分開來。對區別函數二影響力較大的為「錯誤標籤」、「責備他人」與「目擊網路論戰經驗」，能有效區分第二群「迂迴表達網路敵意群體」與第三群的「低網路敵意群體」。

為節省篇幅起見，此處不詳細解說如何以結構方程模式(structural equation modeling, SEM) 建立網路敵意的認知行為模式，但是圖二的認知行為模式經些微修正後，的確是具有適當適配度(index of model fit)的模式，本研究亦曾進行巢狀檢驗(nested test of the model)。本年度之研究結果已分為兩部分，分別投稿到國外及國內期刊中。

民 93 年研究成果：穿梭在現實與虛擬間的攻擊、敵意文章分級系統

「網路論戰與敵意之研究：認知歷程、情境線索與自動化文章敵意分級系統」研究(2/2) 有兩個主要的研究目的：第一是想了解網路語文攻擊行為與其他心理建構的關聯模式研究，把個人在現實世界攻擊相關的特質，如低自尊、社經背景、成人依附等，與網路去抑制環境的特性(如匿名)及網路論戰經驗對個人進行網路攻擊的相對影響力作一比較，以便了解現實世界中個人的攻擊特性對虛擬世界攻擊力的延展影響力。第二個研究目的是建制自動化文章敵意分級系統。

有關研究目的一:網路語文攻擊行為與其他心理建構的關聯模式，所欲探討的研究問題有四：

- 一、男女性在網路論戰經驗頻率、現實生活攻擊行為、網路認知扭曲與網路攻擊四方面的差異情形為何？
- 二、性別、上網時數、上網年資、以電腦為中介的環境特性、現實生活攻擊行為對大學院校學生虛擬網路論戰經驗頻率的相關與預測力為何？
- 三、網路論戰經驗頻率、現實生活攻擊行為對網路認知扭曲的相關與預測力為何？
- 四、上網時數、上網年資、電腦中介溝通環境特性、性別、自尊、內外控人格、依附、社經地位、網路論戰經驗頻率、現實生活攻擊行為、網路認知扭曲對網路攻擊的相關與預測力為何？

本研究正式施測樣本數為 1374 份，取自台灣北部公私立大學共八所，因為網路虛擬世界中並無學校界限，TANET 允許版主轉貼文章，因此無論身處哪一個學校，閱讀相同版面的訊息大致相當。但是私立與公立大學的網路資源相去甚遠，因此本沿就特別著重，讓樣本包含有公私立大學的學生。刪除無效樣本 250 份後(例如：總量表標準差為 0 者、分量表填答方向差異性過大者、完全反向題填答方向卻相反者)，共計

1124 份，詳見表一。在樣本特性方面，男性為 614 人 (54.6%)，女性 510 人 (45.4%)。學歷分佈上，大學部學生 875 名 (77.9%)，碩士班學生 216 名 (19.2%)，博士班學生 33 人 (2.9%)。

表一 正式樣本

學校名	台大	清大	交大	淡江	中華	玄奘	輔仁	文化
原樣本數	277	237	389	82	250	31	73	35
無效樣本	58	40	65	17	45	6	11	8
有效樣本	219 (19.5%)	197 (17.5%)	324 (28.8%)	65 (5.8%)	205 (18.2%)	25 (2.2%)	62 (5.5%)	27 (2.4%)
共計有效問卷為 1124 份								

研究結果如下：

一、男生不管是在網路論戰經驗頻率的總分 ($t=-9.16$, $p<.001$)、目擊網路論戰 ($t=-3.84$, $p<.001$) 以及經驗網路論戰 ($t=-9.95$, $p<.001$) 都比女生來的高，這與黃宏宇 (民 92) 的研究結果相同，可見在現今這網路盛行、男女上網人數已幾乎無差異的時代，目擊與親身參與網路論戰的經驗還是男性較女性為多。

二、網路論戰經驗頻率與性別、一週上網時數、以網路為中介的環境特性、網路次文化、去個人化、去社會抑制、現實生活攻擊行為、身體攻擊、言語攻擊、減敏感性均達顯著正相關。以平均數來看，1124 名受試者平均每週上網時數為 30.16 小時 ($SD=29.70$)，上網年資平均為 5.94 年 ($SD=1.93$)。值得注意的是上網年資與網路論戰經驗總分並無直接的相關。本研究的資料無法說明到底是網路新手—因為不熟悉網路環境與規範因而容易引發論戰，還是網路老手—因為太過熟悉網路文化、享受為所欲為的環境、因而觸犯網規引發論戰，然而研究者推測，此兩種網路使用者均有可能參與論戰，參與論戰的原因不能單純從年資來推斷。

首先進入迴歸公式的是「現實生活攻擊行為」，現實生活中的攻擊可以解釋網路論戰經驗頻率變異量的 14.4% ($F=189.002$, $p<.001$)，標準化回歸係數 β 值為 .293 ($t=9.948$, $p<.001$)，第二個被選入是「電腦中介環境的特性」(以下簡為 CMC 環境特性)，獨力解釋網路論戰經驗頻率變異量的 2.6% ($F=35.113$, $p<.001$)， β 值為 .166 ($t=6.020$, $p<.001$)。第三個被選入的是「性別」，解釋變異量的 1.9% ($F=25.760$, $p<.001$)， β 值為 .134 ($t=4.659$, $p<.001$)，顯示性別是預測網路論戰經驗頻率的重要背景變項，雖然其預測力遠不及於某些心理變項(現實世界攻擊行為)與環境變項(電腦中介環境的特性)。第四個被選入的是「上網年資」，解釋變異量只有 0.5% ($F=6.522$, $p<.05$)。最後被選入的是「上網時數」，解釋變異量的 0.3% ($F=4.100$, $p<.05$)，與上網年資一樣預測力皆不高。

表 4-1-2 網路論戰經驗頻率與各變項間之相關矩陣摘要表

平均數	.55	30.16	5.94	37.00	13.24	8.02	9.36	6.39	26.86	6.19	10.40	10.27	35.71	22.83	12.89
標準差	.50	29.70	1.93	4.72	2.65	1.94	1.75	1.71	7.16	2.11	3.03	3.41	6.53	4.12	4.45
變 項	1	2	3	4	4-1	4-2	4-3	4-4	5	5-1	5-2	5-3	6	6-1	6-2
1.性別	--														
2.一週上網時數	.07*	--													
3.年資	.09**	.05	--												
4.網路中介溝通環境 總分	.11***	-.03	-.03	--											
4-1 網路次文化	.13***	.02	.04	.75***	--										
4-2 去個人化	-.02	-.02	-.10***	.55***	.15***	--									
4-3 社會線索縮減	.03	-.04	.01	.35***	.06*	-.11***	--								
4-4 去社會抑制	.08**	-.02	-.04	.62***	.28***	.26***	-.02	--							
5.現實生活攻擊總分	.38***	.08**	-.06*	.24***	.17***	.24***	-.13***	.25***	--						
5-1 身體攻擊	.22***	.04	-.11***	.16***	.10***	.22***	-.14***	.19***	.85***	--					
5-2 言語攻擊	.30***	.07*	-.06*	.18***	.14***	.18***	-.06	.15***	.80***	.57***	--				
5-3 減敏感性	.32***	.09**	-.01	.24***	.18***	.21***	-.14***	.28***	.86***	.66***	.44***	--			
6 網路論戰經驗總分	.26***	.09**	.06	.25***	.28***	.07*	.01	.16***	.38***	.32***	.26***	.37***	--		
6-1 目擊網路論戰	.11***	.06*	.08**	.23***	.33***	-.06	.09**	.09**	.12***	.03	.16***	.09**	.74***	--	
6-2 經驗網路論戰	.28***	.07*	.01	.15***	.11***	.15***	-.07*	.16***	.45***	.44***	.24***	.46***	.78***	.16***	--

* $p < .05$, ** $p < .01$, *** $p < .001$

表二 以性別、上網時數、上網年資、CMC 環境特性、現實生活攻擊行為總量表預測網路論戰經驗頻率之多元逐步迴歸分析摘要表

投入變項順序	R	R ²	R ² 改變量	F 改變量	B	β	t 值
現實生活攻擊	.380	.144	.144	189.002***	.268	.293	9.948***
CMC 環境特性	.413	.170	.026	35.113***	.230	.166	6.020***
性別	.435	.189	.019	25.760***	1.761	.134	4.659***
上網年資	.440	.194	.005	6.522*	.226	.067	2.464*
上網時數	.443	.197	.003	4.100*	.012	.055	2.025*
常數					17.349		11.114***

* $p < .05$, *** $p < .001$

除性別、上網年資與上網時數外，以「CMC 環境特性量表」、「現實生活攻擊行為」分量表一起進行逐步多元迴歸，以了解各變項分量表對網路論戰經驗頻率的預測力之高低。表三顯示攻擊行為中的「減敏感性」最先進入迴歸模式，可單獨解釋網路論戰經驗頻率變異量的 13.5% ($F=174.669$, $p < .001$)， β 值為 .186 ($t=5.076$, $p < .001$)，表示對於攻擊結果的冷漠與不具同理心能預測網路論戰經驗。

表三 以性別、上網時數、上網年資、CMC 環境特性、現實生活攻擊分量表預測網路論戰經驗頻率之多元逐步迴歸分析摘要表

投入變項順序	R	R ²	R ² 改變量	F 改變量	B	B	t 值
減敏感性	.367	.135	.135	174.669***	.356	.186	5.076***
網路次文化	.427	.182	.047	65.028***	.522	.212	7.836***
性別	.449	.201	.019	26.745***	1.817	.139	4.915***
身體攻擊	.461	.213	.011	15.965***	.468	.151	4.246***
上網年資	.464	.216	.003	4.299*	.188	.056	2.073*
常數					20.145		17.841***

* $p < .05$, *** $p < .001$

三、有關網路論戰經驗頻率、現實生活攻擊行為對網路認知扭曲的相關與預測力，最先進入迴歸公式的是「現實生活攻擊」，可獨力解釋網路認知扭曲的 29.7% 變異量

($F=473.644$, $p < .001$)， β 值為 .527 ($t=19.012$, $p < .001$)。第二個進入迴歸模式的為「網路論戰經驗頻率」，但只能解釋變異量的 1.4% ($F=22.055$, $p < .001$)， β 值為 .142 ($t=5.266$, $p < .001$)，最後進入迴歸模式的為「性別」，可解釋變異量僅佔極小部分 0.9% ($F=15.336$, $p < .001$)， β 值為 -.104 ($t=-3.916$, $p < .001$)。

表四 以性別、網路論戰經驗頻率、現實生活攻擊行為總量表預測網路認知扭曲之多元逐步迴歸分析摘要表

投入變項順序	R	R ²	R ² 改變量	F 改變量	B	β	t 值
現實生活攻擊	.545	.297	.297	473.644***	.425	.527	19.012***
論戰經驗頻率	.557	.310	.014	22.055***	.125	.142	5.266***
性別	.565	.320	.009	15.336***	-.012	-.104	-3.916***
常數					19.021		22.602***

*** $p < .001$

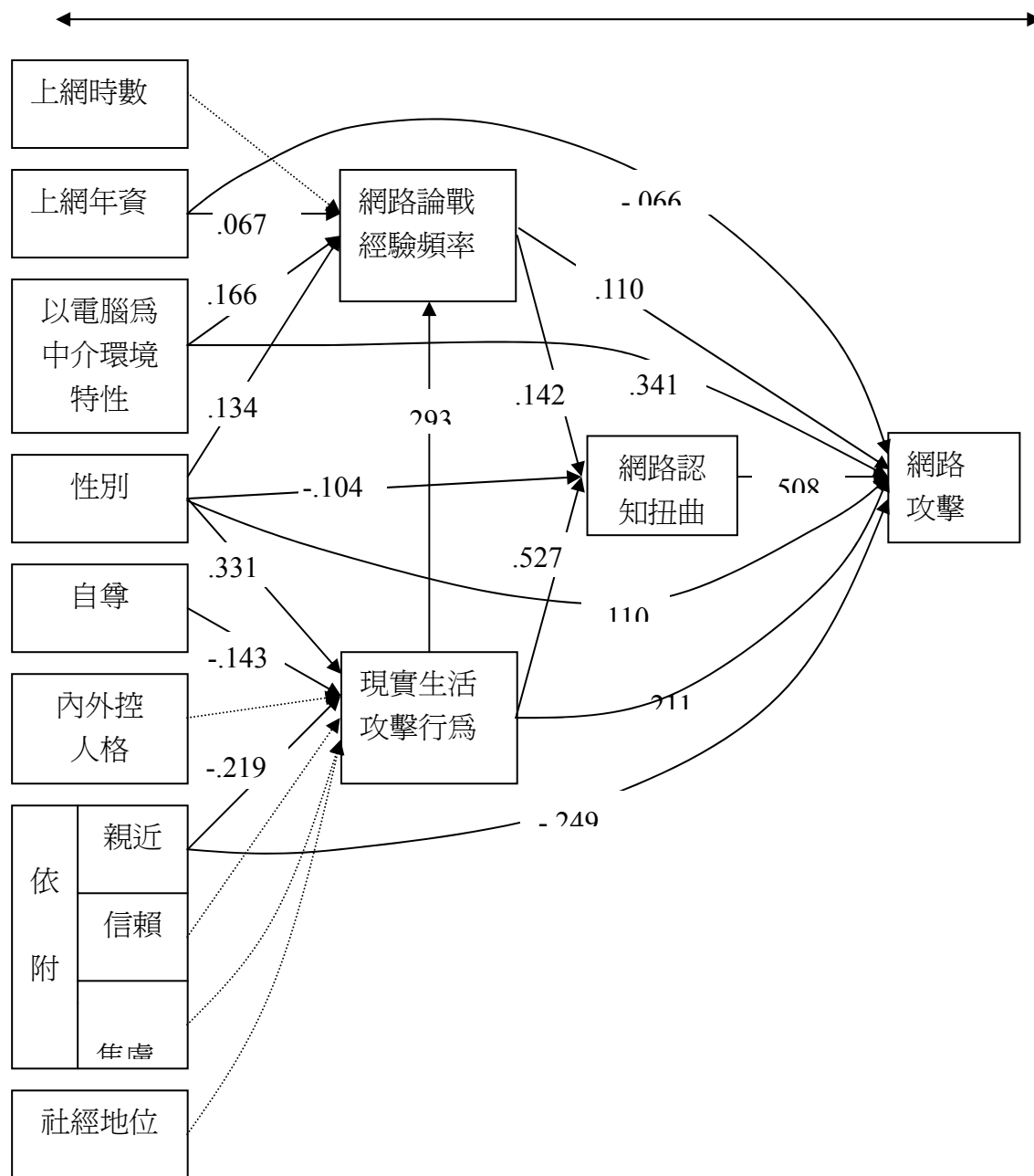
除性別之外，將「現實生活攻擊行為」、「網路論戰經驗頻率」分量表同時丟入，以進一步了解哪一個分量表對網路認知扭曲具較佳預測力。表五顯示，「減敏感性」最先進入迴歸模式，可單獨解釋網路認知扭曲 22.1% 的變異量 ($F = 317.936$, $p < .001$)， β 值為 .280 ($t = 9.318$, $p < .001$)。第二個進入模式的是「言語攻擊」，第三個進入模式的是「經驗網路論戰」，最後進入模式的是「性別」。

表六 以性別、網路論戰經驗頻率、現實生活攻擊行為分量表預測網路認知扭曲之多元逐步迴歸分析摘要表

投入變項順序	R	R ²	R ² 改變量	F 改變量	B	β	t 值
減敏感性	.470	.221	.221	317.936***	.474	.280	9.318***
言語攻擊	.550	.303	.082	131.423***	.640	.336	12.141***
經驗網路論戰	.567	.322	.019	31.490***	.228	.176	6.318 ***
性別	.579	.335	.013	22.132***	-.014	-.125	-.047 ***
常數					20.590		34.108***

*** $p < .001$

四、上網時數、上網年資、電腦中介溝通環境特性、性別、自尊、內外控人格、依附、社經地位、網路論戰經驗頻率、現實生活攻擊行為、網路認知扭曲對網路攻擊的相關與預測力，詳見圖三。



圖三 社會認知理論的網路攻擊整合模式之徑路分析圖(具顯著預測力者為實線，預測力未達顯著者為虛線)

表七 網路攻擊與各自變項間的直接、間接效果分析摘要表

自變項	依變項	直接效果	間接效果	總效果
上網時數	網路攻擊	--	上網時數*網路論戰經驗頻率*網路攻擊=.055*.110=.006	.006
上網年資	網路攻擊	-.066	上網年資*網路論戰經驗頻率*網路攻擊=.067*.110=.007	-.059
CMC 環境特性	網路攻擊	.341	CMC 環境特性*網路論戰經驗頻率*網路攻擊=.166*.110=.018	.359
性別	網路攻擊	.110	性別*網路論戰經驗頻率*網路認知扭曲*網路攻擊+性別*網路認知扭曲*網路攻擊=.134*.142*.508+-.104*.508=-0.043	.067
自尊	網路攻擊	--	自尊*現實生活攻擊行為*網路攻擊=-.143*.211=-.030	-.030
內外控	網路攻擊	--	內外控*現實生活攻擊行為*網路攻擊=.053*.211=.011	.011
親近依附	網路攻擊	-.249	親近依附*現實生活攻擊行為*網路攻擊=-.219*.211=-.046	-.295
信賴依附	網路攻擊	--	信賴依附*現實生活攻擊行為*網路攻擊=-.010*.211=-.002	-.002
焦慮依附	網路攻擊	--	焦慮依附*現實生活攻擊行為*網路攻擊=.004*.211=.0008	.0008
社經地位	網路攻擊	--	社經地位*現實生活攻擊行為*網路攻擊=-.011*.211=-.002	-.002
網路論戰經驗頻率	網路攻擊	.110	網路論戰經驗頻率*網路認知扭曲*網路攻擊=.142*.508=.082	.192
現實生活攻擊行為	網路攻擊	.211	現實生活攻擊行為*網路認知扭曲*網路攻擊=.527*.508=.268	.479
網路認知扭曲	網路攻擊	.508	--	.508

研究目的二：自動化文章敵意分級系統之建立

隨著網路上文章的等比級數增加，如何精確地找出所需要文章成為了重要的議題。本文參酌向量模型應用在自動化文件分類上的研究，提出了針對中文的自動化文章敵意分級系統初探。並從學術網路 BBS 站的硬體討論版(tw.bbs.comp.hardware)抽樣 5000 篇文章，來模擬分級程序，研究發現：

- (1)利用同一主題文章作為訓練文章，來計算敵意與非敵意文章與敵意中心向量的相似度時，其相似度具有明顯差異。
- (2)採用不同主題的訓練文章，來計算同一類文章的相似度亦有差距。
- (3)利用半自動計算，所得到之最佳門檻值0.17來進行分類時，對於非敵意文章有較佳的精確度，其值為0.98，但對於敵意文章的分類精確度則較差，其值為0.25。
- (4)當門檻值降低至0.136時，可得到相等HR 與NHR值，其值為0.73。

【**關鍵詞**】 自動化文件分類、向量模型、資訊檢索、敵意

1. 前言

在網路上公開的論壇或是留言板，由於電腦中介溝通(CMC)的特性，造成網路的無規範狀態[1]，因此常常會看到一些惡意攻擊或是謾罵的文章，這些具有敵意的文章，其內容通常會偏離主題甚多，且容易引起網路上的論戰，對正常使用的討論造成影響，而網管人員也常需要耗費大量時間處理這些文章，因此若能透過自動文件分類的方式，將敵意與非敵意文章分類，則可以幫助網管人員在短時間內篩選出此類發言不當的文章，降低網管人員處理這類文章的頻率，以提昇管理效率。本文的目的，即是希望能以分類敵意文章為例，來探討自動化文章敵意分級系統的程序。

本系統主要功能是對敵意文章進行分類工作，文件分類模型採用G. Salton所提出之向量模型，將測試文章與實際文章以2連字詞的方式進行斷詞後，以Kwok and Luk (2002) [2]針對中文所提出的關鍵詞權重計算公式來計算權重，進而求出敵意文章中心向量(hcv)，最後取得最佳門檻值後，進行抽樣實驗，來檢測分級系統之精確度，並觀察門檻值變化對精確度的影響情形。

2. 本文

本文內容依各小節探討系統相關內容。

2.1 研究樣本文章選取

本研究文章取樣自台灣BBS站的硬體版(tw.bbs.comp.hardware)，抽樣方式採連續抽樣法，樣本文章上傳時間為2003-06-06 04:59:05至2003-06-18 15:38:48。文章長度介於0

與1371字之間。為了避免文章長度影響文章敵意值，因此將針對文章長度進行正規化。實驗前需對訓練文章與實際文章進行敵意相關程度的設定，本實驗對文章是否為敵意文章的相關判定方式，採用次判斷者的判斷方式，本人為判斷者A、另外兩位為次判斷者B與次判斷者C，兩位次判斷者為現職國中資訊組長，對敵意文件的判定都具有相當豐富的經驗，可提高相關認定的一致性與準確度。由於具敵意的文章是由論戰中產生而來，但並不容易瞭解敵意是在論戰的哪一階段產生，因此如要直接判別一篇文章是否為敵意文章難度較高，而論戰文章出現次數較多，較容易觀察，因此本文中對敵意文章的認定，主要是先利用論戰文章篇數出現較多的特性，先計算對於某一主題的文章篇數，若文章篇數明顯高於其他主題的文章篇數，則表示此一主題文章較可能屬於論戰文章，初步定為具有敵意的文章，再利用專家效度進行檢定，從論戰文章中，依照Thompson and Foulger提出的論戰五種過程。分歧〈Divergence〉、爭論〈Disagreement〉、緊張〈Tension〉、敵對〈Antagonism〉、尖銳敵對期〈Profane antagonism〉五個子類別[3]，將文章的敵意程度值設為6~10的整數，數字越高，代表越具有敵意。

2.2 關鍵詞詞庫建立方式

本系統中採用統計式斷詞法來建立關鍵詞詞庫，根據統計發現，在文件中二字詞出現的比率約佔全部詞的75%，在中文文件中，由於較長的詞彙對文件分類並沒有明顯的效果[4]，在Kwok and Luk的研究中發現，在自動化分類中文文章時，若採取向量模式，則斷詞方式採二連字詞，其檢索效能較單字或多連字詞好，因此本系統將以2連字詞為主。而對於只出現一次的二連字詞將不刪除，以提高分類效果。

2.3 敵意文章門檻值

每篇實際文章經過系統計算後，會產生與敵意文章的相似度，此相似度經過正規化之後會介於0與1之間，由於相似度是介於0與1之間的數值，因此必須設定敵意文章的門檻值，才能依門檻值判別是否為敵意文章。為了解以不同主題為訓練文章時的相似度分布情形，我們從12個主題中，選出6個討論篇數較多的主題，也就是文章篇數高於8%的文章，來觀察敵意相似度的分布，此六個主題代碼為A、D、E、F、H、J，並分別以此六主題文章為訓練文章，每次在同一主題中隨機選出三篇文章為訓練文章，再分別從敵意與非敵意文章中，隨機選取30篇文章，觀察相似度的分布情形，研究發現：

- (1)利用同一主題作為訓練文章時，計算出的敵意與非敵意文章相似度，具有明顯差異檢定：敵意文章的門檻值分布在0.2908與0.1089之間，而非敵意文章的門檻值分布在0.1790與0.0853之間，且在所有實驗中，敵意文章的平均相似度都比非敵意文章的平均相似度要高，利用獨立樣本t檢定來進行假設檢定之結果，如表1所示，利用同一主題文章作為訓練文章時，計算出的敵意文章敵意值與非敵意文章敵意值的差異，50%的*** $p < 0.001$ 、50%的* $p < 0.05$ ，達到統計上的顯著水準，可以拒絕虛無假設，也就是當利用同一主題文章作為訓練文件，來計算敵意文章與非敵意文章的相似度時，所計算出的相似度具有明顯差距。這也表示利用同一主題文章所計算出的

hcv，能對敵意與非敵意文章產生明顯的分辨效果。

表1. 不同主題時，敵意文章與非敵意文章敵意值的差異顯著水準

主題A	主題D	主題E	主題F	主題H	主題J
0.002	0.000	0.000	0.048	0.000	0.004

(2)利用不同主題作為訓練文章時，所計算出的敵意文章相似度具有明顯差異：我們利用單因子變異數分析(one-way ANOVA)來進行假設檢定，結果如表2所示：

表2. 以主題為因子之單因子變異數分析摘要

Levene統計	分子自由度	分母自由度	顯著性
3.886	5	174	.002

變異來源	平方和	自由度	平均平方和F	檢定
組間	.35	5	.07	27.7***
組內	.44	174	.02527	
整體	.79	179		

***p<.001

但同質性檢定亦達顯著，無法將此六組資料視為相等，因此我們利用多重比較的LSD法，來觀察各組資料間之相對差異，發現每個主題對於其他主題的差異度皆不同，有些差異甚大，有些組別幾乎沒有差異，但最少都與一個主題有明顯差異，檢定達顯著水準的總數為23個。

(3)利用不同主題作為訓練文章時，計算出的非敵意文章相似度具有明顯差異：利用單因子變異數分析(one-way ANOVA)來進行假設檢定，檢定不同的主題作為訓練文章，所計算敵意文章的相似度是否有明顯差異，結果如表3所示：

表3：以主題作為因子之單因子變異數分析表

Levene統計	分子自由度	分母自由度	顯著性
12.730	5	174	.000

變異來源	平方和	自由度	平均平方和	F檢定
組間	.054	5	.011	4.85***
組內	.389	174	.002	
整體	.44	179		

*** $p < .001$

顯著性*** $p < .001$ ，達非常顯著，但同質性檢定亦達顯著，無法將此六組資料視為相等，因此我們利用多重比較的LSD法，來觀察各組資料間之相對差異，發現每個主題對於其他主題的差異度皆不同，有些差異甚大，有些組別幾乎沒有差異，檢定達顯著水準的總數為14，最少都與一個主題有明顯差異，且具有差異的主題數較敵意文章少。

由以上討論，我們可以知道，利用此六個主題作為訓練文件，能夠使敵意文章與非敵意文章的敵意值，產生明顯差異，而不同主題所計算出之敵意值，亦具有相當差距，顯示每一主題文章在評估敵意時會產生不同影響，所以我們在選擇門檻值的時候，不採取加權方式，也就是以所有能夠判斷正確的敵意與非敵意文章總數，作為門檻值的選擇標準，以同時提高敵意文章與非敵意文章的辨識正確率。

敵意文章與非敵意文章的敵意值大部分皆分布在0~0.3之間，為了設定判別敵意文章與非敵意文章的門檻值，我們將把門檻值從0.04開始，每次以0.01為單位遞增到0.3，來觀察能分辨出敵意文章與非敵意文章的程度，評量標準則是在6個文章類別中，能正確判斷敵意文章與非敵意文章的篇數和。而在門檻值等於0.17時，能正確判別的文章總數和為257，為所有能判別正確的文章總數和的最高值，如圖1所示，因此將以0.17作為敵意文章與非敵意文章的門檻值，相似度大於或等於0.17的文章，判定為敵意文章，而低於0.17的文章，則判定為非敵意文章。

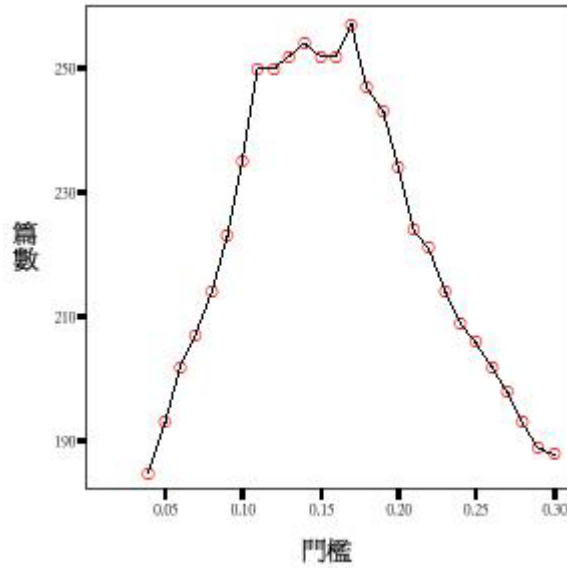


圖1：門檻值與正確判別文章之關係圖

2.4 實驗結果

隨機在訓練文件中選出3篇不同主題文章後，利用5份文章來計算hcv，接下來在5000篇實際文章中，分別對敵意文章與非敵意文章進行實驗，每次隨機抽樣20篇文章，由系統計算出與hcv的相似度。本實驗由於實驗設計的因素，文章分類只有敵意與非敵意兩種，需對評估指標作修正，以符合本實驗所需，評估的指標將使用敵意正確率(HR)與非敵意正確率(NHR)，HR與NHR主要在測量敵意文章與非敵意文章的正确率，數值越高，表示精確度越高。

$$\left\{ \begin{array}{l}
 \text{HR} = \frac{\text{人工與系統皆認為具有敵意的文章篇數}}{\text{抽出文章總篇數}} \\
 \text{NHR} = \frac{\text{人工與系統皆認為不具有敵意的文章篇數}}{\text{抽出文章總篇數}}
 \end{array} \right.$$

研究發現，敵意文章的相似度分布在0到0.213之間，平均值為0.1081，因此大部分文章皆被判別為非敵意文章，NR介0.1與0.9之間，平均值為0.975。敵意文章取樣20篇時，敵意文章的相似度分布在0.0469到0.3194之間，平均值為0.1556，大部分文章皆被判別為非敵意文章，NR介0.1與0.45之間，平均值為0.29。我們發現，在分類非敵意文章時，NHR的值皆平均為0.975，而，因此hcv對於非敵意文章具有良好分類效果，而在分類敵意文章時，平均的準確率只有0.25，可能的原因是敵意文章門檻值過低，而系統在判定敵意文章時，是依據敵意文章門檻值來作為判斷的標準，若門檻值設定過高，會減少人工認

定具敵意文章被系統判定為非敵意文章的篇數，因此準確率就會隨著降低。因此我們試著調整門檻值，來觀察HR與NHR的平均值變化情形。

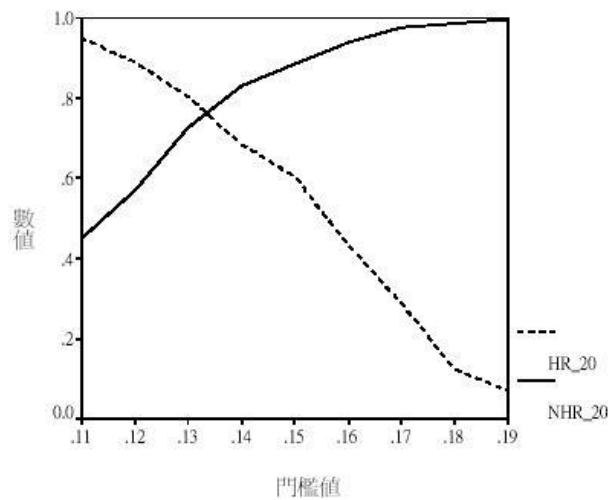


圖2：門檻值與HR、NHR之關係折線圖

當門檻值設定為0.17時，對敵意文章的判別準確率，平均值為0.25，對非敵意文章的判別準確率，平均值為0.97，若降低門檻值，HR與NHR的平均值會同時提昇，當門檻值降低為0.136時，HR會等於與NHR，約為0.7，對於提昇HR與NHR，具有相當明顯的效果。而當取樣篇數為20篇時，亦呈現類似情形。

3. 結論

本文進行了敵意文章分級系統的初探研究，藉由向量模型，將每篇文章對應到一個向量，向量的維度即為hcv的維度，利用餘弦函數計算向量與hcv的餘弦值，此值即為對應文章與hcv之相似度，而相似度高於門檻值的文章稱為敵意文章。並在tw.bbs.comp.hardware中，以連續取樣方式，取樣5000篇文章進行實驗，門檻值的設定方式則依門檻值實驗取得，由人工評定具敵意的文章中，選出討論篇數高於8%的六個主題文章，並分別利用六個主題文章來取得hcv，分別針對敵意與非敵意文章進行相似度計算，利用獨立樣本t檢定與單因子變異數分析來檢測相似度的變化情形，研究發現：

- (1)利用同一主題作為訓練文章時，計算出的敵意文章相似度平均值，皆比非敵意文章相似度平均值要高，且敵意與非敵意文章相似度具有顯著差異，50%的*** $p < 0.001$ 、50%的* $p < 0.05$ ，表示利用同一主題文章所建立的hcv，能明顯的區分出敵意與非敵意文章。
- (2)利用不同主題作為訓練文章時，計算出的敵意文章相似度，具有顯著差異，*** $p < .001$ 。

(3)利用不同主題作為訓練文章時，計算出的非敵意文章相似度，具有顯著差異，
*** $p < .001$ 。

(4)六個主題所計算出的相似度離散差異甚大，有些相似度較集中，有些較分散。

因此在設定門檻值，以六個主題所能正確分辨出的敵意與非敵意文章篇數和為標準，得到最佳的門檻值為0.17，因此以0.17作為正式實驗的門檻值。進行敵意與非敵意文章的分類實驗，並以HR與NHR作為分類效能的指標，發現利用門檻值0.17來進行敵意文章判別，並利用不同主題文章所產生的hcv時，對於非敵意文章有較佳的精確度，約為0.98，但對於敵意文章的分類精確度則較差，約為0.25，但當門檻值調整為0.136時，非敵意文章與敵意文章的精確度皆為0.73，具有良好分類成效。

參考文獻

- [1] Reid, E., Virtual worlds : culture and imagination. From Jones, S. G. (Ed.) ,
Cybersociety : Computer-Mediated Communication and Community, California : Sage
Publications, Inc, 1995.
- [2] Robert W.P. Luk and K.L. Kwok, A Comparison of Chinese Document Indexing
Strategies and Retrieval Models, 2002, ACM Transactions on Asian Language
Information Processing, Vol. 1, No. 3, pp. 225-268, September 2002.
- [3] Thompsen, P. A. & Foulger, D. A., Effects of pictographs and quoting on flaming in
Electronic mail, Computers in Human Behavior, 12(2), pp.225-243, 1996.
- [4] Foo, S. and Li, H., Chinese Word Segmentation Accuracy and It's Effects on
Information Retrieval, Text Technology, 2001.