

## 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)

計畫類別：個別型計畫

計畫編號：NSC 92-2520-S-009-005-

執行期間：2003 年 8 月 1 日至 2004 年 7 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

成果報告類型(依經費核定清單規定繳交)：精簡報告

本成果報告包括以下應繳交之附件：

出席國際學術會議心得報告及發表之論文各一份

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 93 年 5 月 31 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)

計畫主持人：林珊如（交通大學教育研究所）

協同主持人：劉旨峰（中央大學學習與教學研究所）

研究計畫編號：92-2520-S-009-005-

計畫執行起迄：2003.08.01 至 2004.07.31

執行機關：國立交通大學教育研究所

本研究分為兩部分報告，第一部分為網路敵意之概念與測量、相關心理變項的分析，屬於心理學領域的研究成果，這是本年度著重的研究議題。第二部分則為網路敵意文章的自動化分及系統之初探，雖然是明年的研究重點，但是為了建立篩選敵意文章的基本線，今年已著手設計並小規模測試系統。

## 第一部分：網路敵意及其相關因素之研究

### 一、前言與研究目的

根據統計在全球超過四億的上網人口中，台灣經常上網的人口名列全球第 11 名(潘明君，民 90)，上網已逐漸成為台灣的全民運動。隨著資訊科技的發展，與網路人口的增加，改變了過去生活的模式，網際網路的延伸，使得人際關係的互動產生了質變，人們不再需要直接面對面的溝通，而能透過各種不同的網路討論社群，達成意見之交換與情感之滿足(吳妹蓓，民 85；蘇彥豪，民 87)。

但是，由於電腦為中介的傳播模式(Computer-Mediated Communication, CMC)具備匿名的(anonymity)特性，導致許多使用者在失去社會線索的同時，造成自我知覺得降低，產生去個人化(deindividuation)，的情形。在降低自我知覺的情況下，會產生失抑、衝動的行為以及「私刑—暴民」的心態(趙居蓮譯，民 84) Joinson(1998)以社會去抑制(social disinhibition)的觀點來描述網際網路上咒罵與衝突的文字現象。有學者從傳播的觀點，分析網路使用者脫序的文字攻擊行為，於是猥褻性或污辱的言詞就構成所謂的「論戰」(flaming)，不過對於論戰的定義有許多爭議。有學者依據參與論戰者的情緒緊張對峙，由低到高列出層級，每一級均屬論戰範圍(Thompson & Foulger, 1996)，也有學者認為論戰是針對對方的個人攻擊，而非對意見的爭論(Aiken & Waller, 2000)。另外，也有其他學者從網路使用者心理層面進行分析，探討網路衝突的外在現象與內在所產生的網路敵意(Internet hostility)(林珊如、黃宏宇，民 91)。

本研究為了能深入探究網路去抑制的行為，乃決定以大學生為研究對象，針對網路論戰者產生的網路敵意進行分析，修訂現有問卷，在較大樣本群中測試量表之信效

度，並分析影響網路敵意的前置因素，如環境(網路使用時間、經歷網路論戰頻率)、心理病理相關人格(憂鬱、人際焦慮)、人口背景變項，及中介因素(敵意之認知曲解)，如何影響個人於網路環境中呈現出敵意行為，並擬建立網路敵意的社會認知模式。

## 二、論戰

要達到何種外顯的行為才能稱之為論戰 (Flaming, Aiken & Waller, 2000, 國內學者或譯為喧染、爭論、筆戰等，以下通稱為論戰)？研究者的看法莫衷一是 (Kayany, 1998)。文字中的猥褻性 (obscenities) 表達與污辱之言詞，很容易被判定為論戰產生，可是如何釐清激情性 (impassioned) 言語與污辱卻是不容易的。更甚者，有學者認為文句中有「驚嘆」(exclamation) 即視為論戰 (Siegel et al., 1986)。而大部分的學者將論戰視為：針對某個人所進行的攻擊性言詞，其攻擊並非針對此人的論點 (Aiken & Waller, 1999)。

也有學者認為上述定義太廣泛，無法突顯出論戰的內涵與本質 (Lea, O'Shea, Fung, & Spears, 1992)。過去的文獻將論戰具體操作化測量的方式，有以下五種：

- (一) 計算污辱性 (insult) 詞句與指名道姓 (name-calling) 的發生次數 (Siegel et al., 1986; Weisband, 1992)。
- (二) 計算所有社會情緒 (socioemotional) 的表達次數，包括調情 (flirting) 與對他人的情感表達 (Kiesler, Zubrow, Moses, & Geller, 1985)。
- (三) 使用自陳式 (self-report) 量表，讓受訪者陳述溝通時負面的訊息以及論戰發生的場合 (Sproull & Kiesler, 1986)。
- (四) 計算使用「似語言」(para-language) 與咒罵 (incidents of swearing) 的次數，例如語言的非字彙元素 (nonverbal elements of speech, Lea & Spears, 1991)。
- (五) 使用「互動過程分析」(Interaction process analysis, IPA) 測量負面社會情緒內容 (negative socioemotional content, Bales, 1950)。

論戰有諸多的操作性定義，表示眾多學者對構成論戰的本質還沒有共識。更具體的說，論戰之所以為論戰，是因為網路使用者察覺並感受到其負面影響而採取抗爭的行動，此時論戰才產生社會意義 (Thompson & Foulmer, 1996)。不過目前各個論戰的定義都比較著重訊息的知覺 (perception of messages)，也因此產生文字者的主觀佔比較重要的地位，不過並不表示找不到客觀的測量方式。Bale (1950) 發展的互動過程分析，著重於建構主動者與接受者之間引發情緒張力的互動過程。Bale 將負面的社會情緒互動區分為三個層級：不同意 (disagrees) 顯示出緊張 (shows tension) 顯示出敵意 (shows antagonism)，依序表達愈來愈強烈之情緒，這種分類方法廣被許多學者用來定義電腦中介溝通中不同程度的論戰 (Hiltz, Johnson, & Turoff, 1986)。

Thompson 與 Foulmer (1996) 根據 Bale 的分級觀念，利用社會情緒反應 (socioemotional reaction) 中表達緊張 (intensity) 狀態的不同階段，建構出論戰之模式 (a model of flame)。此模式擴大互動分析的層級，增加兩種層級，成為具有過程與階段的模式，用以分析論戰程度的強弱。與互動過程分析模式相較，Thompson 與 Foulmer

不但建構完整的分類系統，且依據電腦中介傳播而設計，其論戰的五等級說明如下：

- (一) 意見紛歧 (Divergence)：在回答問題、爭議、或者是針對某一主題的討論時，出現至少兩種不同的意見。
- (二) 不同意的論點 (Disagreement)：參與者直接提出觀點去反對其他人的論點，但在此階段中並無攻擊他人的觀點，只是提供證據來支持自己的論點。
- (三) 緊張 (Tension)：參與者不但提出對立的觀點，且攻擊他人的論點。反駁對方的質疑論點，並不斷膨脹自己論點的正確性。
- (四) 敵意產生 (Antagonism)：參與者不但攻擊對立的論點，也攻擊對立的個人。透過指名道姓 (name-calling) 與針對個人特性 (ad hominem) 的攻擊，損害對手的人格與信用，而較少關注於原先兩人爭論的焦點。
- (五) 褻瀆式的恐嚇 (Profane antagonism)：參與者彼此間充滿著過度的敵意，互相攻擊，使用褻瀆的語言、自負的攻擊性文字、與低劣的攻擊性論點，此時已完全忽略最初意見紛歧的論點為何。

而 Lea 與 Spears (1991) 整理過去的文獻發現在電腦中介溝通的去抑制行為與論戰可能來自以下五個因素：缺少社會線索、去個人化、人際協調與回饋較困難、注意過度集中、以及順從網路次文化的規範。

不過，有些訊息對某些人來說是侮辱或攻擊，對其他人而言則為平常，因此認知解釋的歷程或個人主觀成分在分析論戰時便是一個重要的中介變項，就如同上述，論戰的產生必須個體察覺並採取相對應的反應時，才具有意義存在 (Thompson & Fouler, 1996)。

由於網路論戰的分類是針對網路使用者的文字描述進行分析，是溝通完成後的結果之探討，並無法得知網路使用者是屬於何種心理狀態，導致產生網路論戰的發生。透過內容的分析，可以發現網路上的論戰行為是屬於敵意層面，也就是網路使用者在不滿其他人或相關情境下所做出的攻擊行為。因此，為了深入探討網路使用者的敵意狀態，與設計出具有良好信效的的測量工具，則必須針對過去敵意相關研究加以整理。

### 三、敵意之相關研究

敵意 (hostility) 與攻擊 (aggression) 的概念有很高的重疊性 (李建德, 民 89)。部分學者認為敵意的內涵較為廣泛，包括人格與行為層面，例如生氣、憤怒、負向行為、直接或間接攻擊等 (Megargee, 1985)，因此敵意包含認知與行動層面。但也有學者將敵意視為攻擊的認知成分，敵意促發生氣的情緒，導致攻擊行為發生 (Spielberger, Reheiser, & Sydeman, 1995)。不論如何，敵意與攻擊間的關係是密切的。

敵意是一複雜的概念，不同的學者對敵意有不同的定義與測量方式，因此往往研究結果不盡相同 (Meesters, Muris, & Backus, 1996)。舉例而言，過去的敵意研究常將焦點放置在其與冠狀動脈心臟病 (coronary heart disease, CHD) 的相關，結果發現敵意是促發冠狀動脈心臟病的主要因子之一 (Barefoot, Dodge, Peterson, Dahlstrom, &

Williams, 1989; Siegman, Dembroski, & Ringel, 1987) 但是也有部分研究反駁這樣的觀點, 認為敵意與冠狀動脈心臟病並無直接的關連 (Leon, Finn, Murray, & Bailey, 1988; Hearn, Murray, & Luepker, 1989)。

除了心血管疾病的研究外, A 型的人格特質也與敵意有正向關聯 (Friedman & Rosenman, 1974)。A 型人格的特色為沒有耐心、注重時間精確、愛好競爭、情緒起伏不定且要求完美 (Meesters, Muris, & Backus, 1996)。而從一些敵意量表, 如 MMPI、Cook-medley 與 HO 量表, 皆發現量表分數過高, 則其心血管疾病、死亡率, 與癌症罹患率均增高 (Smith, 1992)。由此可知, 過去敵意之研究, 目的在於瞭解個體內在心理狀態, 並能預測往後的生理狀態。

有學者視敵意為攻擊的信念, 也有學者視敵意為人格特質, 也有學者廣義地將敵意包含認知情感與行為三個層面。這三種觀點分述如下:

- (一) 敵意僅是攻擊前的認知歷程: Buss 與 Perry (1992) 認為敵意僅是攻擊前的認知歷程, 因此敵意並無行動之層面, 僅是對於人或事物的負面評價。
- (二) 敵意是穩定的人格特質: 部分學者認為敵意是一種穩定的人格特質, 當個體遭受不公平待遇時, 會形成一套負面評價與認知歷程的準則, 不易隨情境改變而變動 (Eckhardt & Deffenbacher, 1995; Huesmann, Eron, Lefkowitz, & Walder, 1984; Megargee, 1985)。
- (三) 敵意同時涵蓋認知情感與行為層面: 例如 Barefoot, Dodge, Peterson, Dahlstrom 與 Williams (1989) 認為敵意是一廣泛的概念, 是個體在人際互動中所形成的負向認知、情緒與行為。Smith (1992) 也指出敵意是一組負向態度、信念與評價, 因此敵意具備了情感、認知與行為三層面。

本研究擬以廣義的認知行為歷程來定義敵意, 如同 Barefoot 等人 (1989) 與 Smith (1992) 對於敵意之界定: 個體若察覺自身遭受不公平或威脅之對待時, 伴隨負面情緒的激發, 導致負面評價的認知歷程, 進而採取負面的因應行為, 例如言語攻擊、生理攻擊或間接攻擊等等。所以敵意同時涵蓋三種歷程, 即情感、認知與行為。

#### 四、敵意與攻擊之測量

對於敵意的測量方面, 不同的立論取向影響到採用何種測量工具。以 Cook-Medley Hostility Inventory 而言, 該量表為二點量表, 自從 1957 年發展至今, 內部一致性係數與再測信度能達到 .80 以上 (Barefoot et al., 1983; Cook & Medley, 1954; Smith, 1992)。但是在效度方面卻顯示出種種不同的因素結構 (Barefoot et al., 1983)。

Buss 與 Durkee 發展的 Buss-Durkee Hostility Inventory (BDHI), 以七個向度來測量敵意: 攻擊 (assault) 間接攻擊 (indirect aggression) 言語攻擊 (verbal aggression) 憤怒 (irritability) 負向態度 (negativism) 怨恨 (resentment)。然而經過不同學者進行因素分析, 各有不同的因素結構。例如 Buss 與 Durkee (1957) 歸納出兩個因素:

攻擊 (aggression) 與敵意 (hostility)。Bendig (1962) 也發現兩個因素：明顯 (overt) 與隱藏 (covert) 的敵意。此外，Siegman、Dembroski 與 Ringel (1987) 則整理出表現 (expressive) 與神經質 (neurotic) 的敵意。最後，Bushman、Cooper 與 Lemke (1991) 則對各種不同的因素分析進行後設分析，證實了兩個因素的結構是合理的。

Buss 與 Perry (1992) 依據 BDHI 修正成為 Aggression Questionnaire (AQ)，為李克特氏五點量表，起初包含了六個敵意向度：身體的攻擊 (physical aggression)、語言的攻擊 (verbal aggression)、生氣 (anger)、間接攻擊 (indirect aggression)、怨恨 (resentment)、懷疑 (suspicion)。經過一系列的因素分析，最後修改為四個向度，共 29 題，其向度包含：身體的攻擊、語言的攻擊、生氣、敵意。身體攻擊與語言攻擊之定義為實際的攻擊他人；生氣的定義是「生理的喚起，並預備進行攻擊」；而敵意的定義是「感受到不好的意圖與不公平的事件，且大多屬於攻擊的認知成分」。雖然此量表有不錯的信效度，但是向度的多寡與命名，還存有爭議。四種因素的結構，在大學生為受試樣本的研究中獲得支持 (Harris, 1995; Meesters, Muris, Bosma, Schouten, & Beuving, 1996)；但在罪犯的研究中，卻發現兩種因素的敵意結構較為符合 (Williams, Boyd, Cascardi, & Poythress, 1996)，也就是身體攻擊與生氣結合為一因素，而語言攻擊與敵意結合為另一因素。

雖然學者對於敵意與攻擊的建構內涵意義有不一致的觀點，但是對於敵意與攻擊之區分卻有相當的共識。大部分學者都同意敵意是屬於認知成分居多，而攻擊多屬於外顯行為之表現。Buss 與 Perry (1992) 認為在 AQ 中，「生氣」能有效地連結屬於認知部分的「敵意」，與實際行為成分的「身體攻擊」和「語言攻擊」。而且敵意是促發生氣的主要原因，並引發下一步的攻擊行為 (Spielberger, Reheiser, & Sydeman, 1995)。

國內也發展出一些敵意或攻擊的心理測驗 (如：柯永河、林幸台、張小鳳，1993；林一真，1995)。柯永河等 (1993) 根據 Mauger 與 Adkinson 編制成適合國人的入際行為調查表 (Interpersonal Behavior Survey, IBS)，其中包括四個分量表，分別為：效度量尺、攻擊量尺、自我肯定量尺，與關係量尺。在攻擊量尺中，包含了六個因素：

- (一) 敵對 (Hostility)：測量對他人的攻擊傾向，視攻擊為超越別人或保護自己的正當理由。
- (二) 表現忿怒 (Expression of Anger)：易發脾氣、強力且直接表示其憤怒的傾向。
- (三) 忽視權利 (Disregard for Rights)：藉忽視他人權利，以保護自己或獲取利益。
- (四) 口頭攻擊 (Verbal Aggressiveness)：以言語批評、攻擊、貶抑他人。
- (五) 身體攻擊 (Physical Aggressiveness)：以實際或想像的肢體動作攻擊他人。
- (六) 消極攻擊 (Passive Aggressiveness, PA)：以間接或消極的方式攻擊他人的行為傾向，如抱怨、拖延、頑固等。

本研究將參考柯永河等 (1993) 的「攻擊量尺」與 Buss & Perry (1992) 的 Aggression Questionnaire 中的向度，加以修正成適合於網路情境的攻擊敵意試題，將認知與行動的部分皆納入測驗的範圍。

## 五、研究問題

1. 網路敵意認知行為模式之遠端、近端變項相關考驗皆達顯著。
2. 高低網路敵意群體在認知行為模式中遠端與近端變項之差異性考驗皆達顯著。
3. 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。

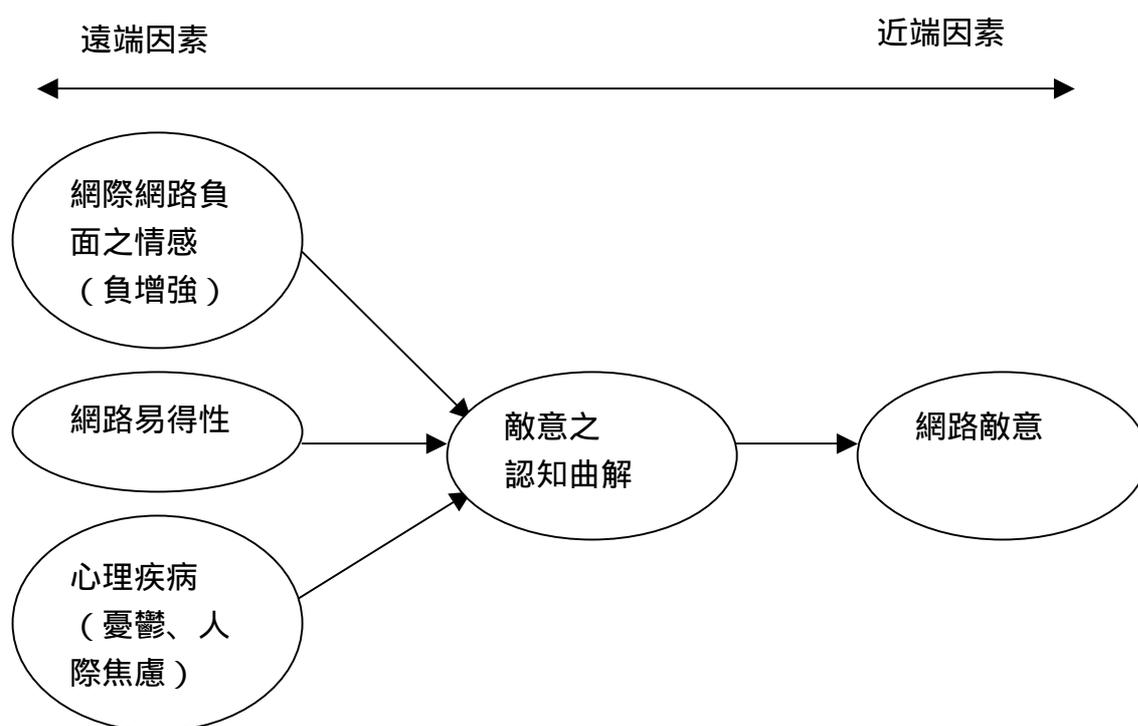


圖 1-1 網路敵意之認知行為模式

## 六、研究方法

### (一) 樣本

本研究預試以國立交通大學學生為預試對象，於在民國 91 年 12 月 20 日至 92 年 1 月 5 日期間，在交通大學資訊科學 BBS 站校園討論區中徵求預試受試者，並進行紙本版的預試問卷，受試群體包含大學生與研究生。有效樣本為 156 份，回收資料後加以分析整理，修改問卷內容後，進行正式施測。

正式施測，在參考各校 BBS 站流量的上網人數後，決定以兩所公立學校：台灣大學、交通大學，與一所私立學校：淡江大學為研究樣本，並且以網路問卷的形式實施。網址為 <http://sunnylin74.cte.nctu.edu.tw/question/>，網頁開放時間為民國 92 年 3 月 24 日至 3 月 31 日期間。正式施測的樣本包含，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% ) 。

## (二)研究工具

本研究用來收集資料的工具有下列八種：

### 1. 網際網路情境線索量表 ( Questionnaire of Internet Situational Cues, QISC )

概念型定義為接觸網際網路或新科技的經驗，對於個體可能會成為增強經驗。即個體上網時，若從使用網路獲得正向的回應，根據操作制約原則，該刺激反應會再度加強原有之行為。如果刺激之間產生連結，則會透過次級增強的原則，提供更多情境上的線索來增強病態網路使用的行為 ( Davis, 2001 ) 。 增強經驗之測量採自編「網際網路情境線索」量表，分成兩個部分，分別為初級增強與次級增強，共 11 題，採李克特式四點量表。以該量表的總分測量網路使用者情境線索增強之程度。

### 2. 網路論戰經驗頻率量表 ( Online Flaming Experience, OFE )

概念型定義為個體在網路上若察覺自身遭受不公平或威脅之對待時，伴隨負面情緒的激發，導致負面評價的認知歷程，進而採取負面的因應行為，例如論戰的文字表現，直接或間接的口語攻擊等等。本部分參考「網路衝突經驗頻率量表」( 林珊如、黃宏宇，民 91 ) ，與論戰之分類 ( Thompsen & Fouler, 1996 ) ，改編成為「網路論戰經驗頻率量表」，分成兩個向度，共 16 題，採李克特式四點量表。分別為目擊網路敵意事件，與經驗網路敵意事件，而網路敵意事件則依據 Thompsen 與 Fouler ( 1996 ) 之分類，採用論戰中敵意程度較高的「緊張」、「敵意產生」與「褻瀆式恐嚇」三項加以設計，並以該量表總分測量網路使用者在網路上負面情感之頻率。

### 3. 網路易得性 ( Time of Internet Use, TIU )

概念性定義為個體接觸網路的便利性與可得性，藉由網路的易得性，個體可長期使用網際網路。本研究計算每天上網時數，與每週上網時數，即網路易得性。分數愈高，上網時間也就愈長。

### 4. 憂鬱 ( Depression )

概念性定義乃依據 Beck 的認知理論，具有憂鬱傾向的患者對於自身有強烈的負面的不合理看法，認為自己是無價值、前途黯淡，情緒非常沮喪，並發展形成負面的認知基模 ( 引自洪蘭，民 84 ) 。 以「貝克憂鬱量表第二版」( The Beck Depression Inventory-Second Edition, 以下簡稱寫為 BDI- ，陳心怡譯，民 89 ) 測量受試者的憂鬱程度。

### 5. 人際焦慮 ( Interpersonal Anxiety )

概念性定義為焦慮反應起於害怕某些可怕的事物要發生在自己身上，於是引起嚴重的緊張，使個體無法正常工作，並預期災難即將來臨 ( 引自洪蘭，民 84 ) 。 而個體

在人際關係中出現焦慮感受的程度，例如害怕被拋棄，或擔心不被他人所喜愛，即為人際焦慮。採用 Collins 與 Read( 1990 )所編制「成人依附量表」( Adult Attachment Scale ) 中的「人際互動焦慮程度」分量表，以該量表總分測量。

#### 6. 敵意的認知曲解 ( Hostile Cognitive Distortions )

概念性定義為是一種不真實、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 ( 鄭夙雅，民 90 )。參考連秀鸞 ( 民 88 ) 所編制的「認知扭曲量表」，並進行修正，問卷共分成四個分量表：為負向思考、錯誤標籤、自我中心、責備他人。得分愈高者，其認知曲解程度愈高，反之亦然。

#### 7. 網路敵意 ( Internet Hostility )：

概念性定義為個體在網路上若察覺自身遭受不公平或威脅之對待時，伴隨負面情緒的激發，導致負面評價的認知歷程，進而採取負面的因應行為，例如論戰的文字表現，直接或間接的口語攻擊等等。本研究採用林珊如、黃宏宇 ( 民 91 ) 編制的「網路敵意問卷」( Internet Hostility Questionnaire, IHQ )，問卷有六個分量表：網路人際懷疑、表現忿怒、忽視權利、直接口語攻擊、運用團體力量攻擊、其他網路攻擊手段。以總分測量網路敵意之程度為何。

## 七、研究結果

### (一) 網路敵意認知行為模式之遠端、近端變項的關聯性

為了回答研究問題 1：「網路敵意認知行為模式之遠端、近端變項相關考驗皆達顯著」，進行下列之考驗。從表 1-1 發現，網路敵意總分與敵意的認知曲解總分呈現高相關 ( $r = .63, p < .001$ )，而網路敵意各向度也與敵意認知曲解各向度達到正相關

( $r = .09- .58, p < .01- .001$ )，由此可以觀察出敵意方面的不合理認知，的確會影響在網路上的敵意表現。在網路論戰經驗頻率得分上，經歷網路論戰頻率愈高者，在網路敵意表現上愈高 ( $r = .47, p < .001$ )，而經歷網路論戰經驗與網路敵意之相關 ( $r = .53, p < .001$ )，與目擊網路論戰經驗與網路敵意之相關 ( $r = .16, p < .001$ ) 皆達顯著。各分向度的分析上，除了網路人際懷疑與目擊網路論戰頻率，其他攻擊手段與目擊網路論戰之相關，忽視權利與經歷網路論戰頻率之間相關未達顯著外，其他各向度間之相關皆達顯著水準 ( $r = .12- .61, p < .001$ )。目擊網路論戰的經驗是個人以旁觀的角度，看到網路上眾人互相謾罵，並不造成實質上的傷害，或許因此與網路的不信任感之間的關係較為薄弱；同樣地，其他網路攻擊手段也是類似的情況。至於經歷網路論戰與忽視權利之間關係的不明顯，可能因為經歷過網路論戰後造成個體實質損害，較容易採取其他直接的方式來表達不滿 (例如直接口語攻擊，相關值為 .45)，而捨棄使用間接的方式表達不滿。

在人際焦慮方面，分數愈高者，網路敵意愈高 ( $r = .20, p < .001$ )。網路敵意也與憂鬱呈現正相關 ( $r = .16, p < .001$ )。值得注意的是，除了直接口語攻擊 ( $r = .10- .13, p < .01- .001$ ) 其他網路攻擊手段與網路使用時間呈現正相關外 ( $r = .08- .09, p < .01$ )，其餘各變敵意項皆與使用者時間未達到顯著水準，可能原因來自於直接口語攻擊與其他網路攻擊手段是屬於外顯且直接的表達方式，相對於其他敵意表達方式而言，網路使用者若取採直接行動將會耗時在網路上，而其他表達敵意方式則較無時間上的差異。在林珊如、黃宏宇 (民 91) 的先前研究中發現，上網時間與網路敵意之表現無必然之關係，但上網時間與網路敵意之間存在中介變項，透過使用網路後所感受到主觀的負面感受之調節作用，間接影響網路敵意之展現。

## (二) 高低網路敵意群體在認知行為模式中遠端與近端變項之差異性

為了回答研究問題 2：「高低網路敵意群體在認知行為模式中遠端與近端變項之差異性考驗皆達顯著」，進行下列之考驗。研究者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。

首先，針對敵意認知曲解上之差異進行分析，從表 1-2 多變量變異數分析可以發現，高低網路敵意群體在敵意認知曲解整體上具有顯著差異 ( $Wilk's \lambda = .65, p < .001$ )。而在各向度與總分之  $t$  考驗中也同樣發現 (見表 3)，高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 ( $t = -19.04, p < .001$ ) 上，皆較低網路敵意群體來得高。據此可以支持認知行為模式不但符合現實世界中的敵意攻擊，也可以應用於虛擬網路世界中論戰的心理狀態。

表 1-2 網路敵意高低群體在敵意認知曲解之多變量變異數分析摘要表

變異來源	Df	SSCP				Wilk's
組間	1	1566.82				.65***
		1525.12	1484.53			
		2116.31	2059.99	2858.51		
		1030.26	1002.84	1391.57	677.44	
組內	670	6181.00				
		2414.71	4990.87			
		2906.30	2532.10	6937.20		
		1009.23	1564.01	2098.47	2756.84	

\*\*\* $p < .001$



表 1-1 網路敵意與各變項積差相關摘要表

平均數	8.06	9.64	13.45	7.11	12.38	8.09	58.73	19.87	12.00	31.86	9.76	11.47	10.97	6.61	38.81	14.46	30.80	4.32	30.82
標準差	2.31	2.93	3.17	2.53	2.91	3.41	10.46	4.38	4.94	6.99	3.18	2.96	3.58	2.16	9.44	4.04	8.09	2.91	21.89
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1.網路人際懷疑	--																		
2.表現忿怒	.33***	--																	
3.忽視權利	.19***	.04	--																
4.直接口語攻擊	.22***	.52***	-.04	--															
5.運用團體力量攻擊	.06	.29***	.21***	.40***	--														
6.其他網路攻擊手段	.34***	.46***	-.05	.57***	.13***	--													
7.網路敵意總分	.55***	.72***	.39***	.72***	.57***	.69***	--												
8.目擊網路論戰	.04	.12***	.16***	.13***	.16***	-.02	.16***	--											
9.經歷網路論戰	.30***	.45***	-.04	.45***	.14***	.61***	.53***	.12***	--										
10.網路論戰經驗總分	.23***	.39***	.07*	.40***	.20***	.42***	.47***	.71***	.78***	--									
11.負向思考	.35***	.36***	.15***	.36***	.20***	.39***	.49***	.10**	.40***	.34***	--								
12.錯誤標籤	.16***	.38***	.19***	.37***	.42***	.20***	.47***	.12***	.21***	.22***	.51***	--							
13.自我中心	.26***	.42***	.11***	.51***	.27***	.51***	.58***	.12***	.21***	.22***	.53***	.54***	--						
14.責備他人	.09**	.34***	.13***	.41***	.40***	.25***	.45***	.27***	.09**	.25***	.23***	.51***	.54***	--					
15.敵意認知曲解總分	.29***	.48***	.18***	.53***	.39***	.45***	.63***	.13***	.41***	.37***	.78***	.81***	.85***	.71***	--				
16.人際焦慮	.02	.20***	-.02	.19***	.16***	.15***	.20***	.01	.15***	.11***	.19***	.21***	.22***	.19***	.26***	--			
17.憂鬱	.05	.15***	.02	.14***	.10**	.11***	.16***	.11**	.07*	.12**	.18***	.18***	.21***	.17***	.24***	.27***	--		
18.每天上網時間	-.03	-.03	-.01	.10**	.03	.08**	.04	.11***	.09**	.13***	.07*	.04	.09**	.02	.08*	.07*	.08**	--	
19.每週上網時間	-.01	-.01	-.01	.13***	.02	.09	.06	.11**	.09**	.13***	.09**	.05	.10***	.03	.09**	.07*	.09**	.92***	--

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

表 1-3 網路敵意高低群體在敵意認知曲解之 t 考驗摘要表

變項	網路敵意	個數	平均數	標準差	t 值
負向思考	低網路敵意	342	8.41	2.81	-13.00***
	高網路敵意	330	11.47	3.26	
錯誤標籤	低網路敵意	342	9.93	2.81	-14.12***
	高網路敵意	330	12.90	2.64	
自我中心	低網路敵意	342	9.12	2.79	-16.54***
	高網路敵意	330	13.25	3.61	
責備他人	低網路敵意	342	5.56	2.06	-12.83***
	高網路敵意	330	7.57	1.99	
敵意認知曲解總分	低網路敵意	342	33.02	7.86	-19.04***
	高網路敵意	330	45.18	8.69	

\*\*\* $p < .001$

表 1-4 呈現高低網路敵意群體在網路論戰經驗頻率上之整體差異，多變量變異數分析達顯著水準 (Wilk's  $\lambda = .76, p < .001$ )，而在各向度與總分的考驗上 (見表 1-5)，高網路敵意群體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 ( $t = -13.29, p < .001$ ) 上，皆較低網路敵意群體來得高。在林珊如與黃宏宇 (民 91) 的研究中也有類似的發現：網路衝突經驗頻率與網路敵意間具有正向路徑關係存在。因此可以推論，網路論戰經驗的確會影響主觀的網路敵意之表現，但其路徑關係則有待往後章節再以其他統計分析加以確認。

表 1-4 網路敵意高低群體在網路論戰經驗之多變量變異數分析摘要表

變異來源	df	SSCP		Wilk's
組間	1	488.83		.76***
		1419.47	4121.87	
組內	670	12668.49		
		271.19	14782.12	

\*\*\* $p < .001$

表 1-5 網路敵意高低群體在網路論戰經驗之 t 考驗摘要表

變項	網路敵意分組	個數	平均數	標準差	t 值
目擊網路論戰	低網路敵意	342	19.00	4.23	-5.09***
	高網路敵意	330	20.71	4.46	
經歷網路論戰	低網路敵意	342	9.90	2.84	-13.51***
	高網路敵意	330	14.86	6.05	
網路論戰經驗頻率總分	低網路敵意	342	28.90	5.61	-13.29***
	高網路敵意	330	35.56	7.24	

\*\*\* $p < .001$

網路敵意的高低群體在憂鬱、人際焦慮與網路使用時間的差異性考驗上，由表 1-6 可以發現，高網路敵意的群體在貝克憂鬱量表 ( $t=-3.96, p<.001$ )、人際焦慮 ( $t=-5.44, p<.001$ ) 上，較低網路敵意群體來得高。但是在網路使用時間上，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上的關係並不明顯，亦即屬於「量」的上網時間，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，質會比量來的重要。

表 1-6 網路敵意高低群體在憂鬱、人際焦慮、網路使用時間之 t 考驗摘要表

變項	網路敵意分組	個數	平均數	標準差	T 值
貝克憂鬱量表	低網路敵意	342	29.93	7.39	-3.96***
	高網路敵意	342	32.45	9.00	
人際焦慮	低網路敵意	342	13.81	4.31	-5.44***
	高網路敵意	330	15.50	3.75	
每天上網時間	低網路敵意	342	4.31	2.77	-0.74
	高網路敵意	330	4.48	2.88	
每週上網時間	低網路敵意	342	30.57	20.54	-1.04
	高網路敵意	330	32.30	22.50	

\*\*\* $p<.001$

### (三)網路敵意之群集分析

為回答研究問題 3：「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？」因此，研究者將蒐集之全部樣本進行群集分析，但是研究樣本過多 (987 名)，統計軟體在處理資料上耗時且不易完成 (張紹勳、張紹評、林秀娟，民 82)，所以研究者以二階段法進行資料的分類。首先，以隨機取樣的方式，從全部樣本中抽取 188 名受試樣本 (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區分群體，結果 188 名樣本共可區分成三群體 (見表 1-7)。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。而在事後比較與各平均數比較後，研究者將這三群加以命名。

第一群在「忽視權利」與「運用團體力量攻擊」上屬於高分群，參考該類型題目之後，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，因此將此群集命名為「迂迴表達網路敵意群體」，共 92 名。而第二群在「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，而且是屬於直接、衝動表達敵意傾向，因此命名為「直接表達網路敵意群體」，共 52 名。第三群體在網路敵意各向度上皆是屬於低分群體，網路敵意表現較不明顯，因此研究者將此群體命名為「網路論戰沈默群體」，包含 44 名。

表 1-7 網路敵意集群分析摘要表 ( 隨機樣本 188 份 )

因素別	群 集			F 值
	迂迴表達 網路敵意群體	直接表達 網路敵意群體	網路論戰 沈默群體	
網路人際懷疑	8.04	<b>9.18</b>	7.00	7.41**
表現忿怒	9.29	<b>12.41</b>	7.59	57.41***
忽視權利	<b>15.43</b>	12.83	10.55	55.13***
直接口語攻擊	6.56	<b>9.84</b>	5.46	69.61***
運用團體力量攻擊	<b>13.10</b>	<b>13.33</b>	10.19	18.81***
其他網路攻擊手段	6.60	<b>12.40</b>	6.57	139.45***
群體個數	92 ( 49% )	52 ( 28% )	44 ( 23% )	

\*\* $p < .01$  \*\*\* $p < .001$

迨隨機樣本分析完成後，從資料中整理出各群體的中心點，並開始進行第二階段的分群。第二階段採用非階層集群分析，並訂出三群的中心進行分類，分群結果與效度檢驗列於表1-8。可以觀察出，分群的結果符合預期，也與先前隨機樣本的結果相同，顯示在網路世界中論戰的群體，的確有敵意表達強度與敵意表達方式上的差異。

表1-8 全部樣本之網路敵意集群分析摘要表 ( 全部樣本987份 )

因素別	群 集			F 值
	迂迴表達 網路敵意群體	直接表達 網路敵意群體	網路論戰 沈默群體	
網路人際懷疑	8.15	<b>9.41</b>	6.91	65.68***
表現忿怒	9.39	<b>12.63</b>	7.66	287.36***
忽視權利	<b>15.51</b>	13.09	10.09	355.41***
直接口語攻擊	6.66	<b>10.02</b>	5.11	379.58***
運用團體力量攻擊	<b>12.78</b>	<b>13.26</b>	9.38	128.83***
其他網路攻擊手段	6.48	<b>12.98</b>	6.30	620.29***
群體個數	468 ( 47% )	255 ( 26% )	264 ( 27% )	

\*\*\* $p < .001$

為了瞭解網路敵意不同的群體之影響因素為何，嘗試由病態網路使用相關各變項、網路敵意相關各變項、網路使用之背景變項等，將網路敵意群體的區分意義顯示出來，因此進行區別分析(discriminant analysis)。

表1-9顯示區別分析產生兩組區別函數，均達到顯著水準( Wilk's 值分別為 .575 與 .896,  $p < .001$  )。以區別函數一而言，區別分數與類別依變項的關連強度為 .599 (  $p < .001$  )，而區別函數二的區別分數與類別依變項的關連強度則為 .322(  $p < .001$  )。從結構矩陣的區別負載係數可以觀察各變數對於此函數影響力的多寡。一般而言，若

係數的絕對值大於 .30，表示自變數具有區別能力；而係數大於 .45時，則代表該變數具有「強」的區別能力（張紹勳、張紹評、林秀娟，民82）。

在區別函數一中，「經歷網路論戰經驗」、「自我中心」、「負向思考」、「錯誤標籤」、「病態網路使用認知曲解」、「責備他人」、「病態網路使用」具有不錯的區別能力。而且圖1-2顯示出三群體在區別函數上之分佈，區別函數一可以區分出「直接表達網路敵意群體」與「迂迴表達網路敵意群體」。如果欲區分「迂迴表達網路敵意群體」與「網路論戰沈默群體」則以函數二較為合適，其變項對於此函數影響較大的為「錯誤標籤」、「責備他人」與「目擊網路論戰經驗」。

「直接表達網路敵意群體」的特性為較常親身參與網路論戰，且其認知曲解程度也較迂迴表達網路敵意群體來得高。不過值得注意的是，在病態網路使用上，愈高的病態網路使用程度較能區分與預測迂迴表達網路敵意群體。因此，在網路上表達直接且衝動的敵意使用者，與病態網路使用並無絕對之關係，相反地，在人數最多的迂迴表達網路敵意群體中，病態網路使用較具有預測力。

表1-9 區別分析摘要表

變數名稱	區別函數一		區別函數二		平均數		
	區別函數係數	區別負載係數	區別函數係數	區別負載係數	迂迴表達網路敵意群體	直接表達網路敵意群體	網路論戰沈默群體
Wilk's 典型相關( )	.575***	.599	.896***	.322			
經歷網路論戰	.658	.811++	-.396	-2.83	10.81	16.33	9.92
目擊網路論戰	-.007	.102	.460	.416++	20.38	20.09	18.73
自我中心	.387	.697++	-.239	.110	10.58	13.58	9.13
負向思考	.091	.534++	-.005	.185	9.59	11.57	8.31
錯誤標籤	.129	.459++	.788	.712++	11.85	12.65	9.65
責備他人	.070	.438+	.194	.370+	6.67	7.56	5.60
病態網路使用認知曲解	.117	.449+	.262	.100	33.29	37.36	31.11
人際焦慮	.000	.205	-.108	-0.064	14.19	15.50	13.95
憂鬱	.007	.151	-.274	-0.124	30.19	32.45	30.33
病態網路使用	-.024	.321+	-.092	-0.002	47.24	52.6	45.27
網路情境線索	.056	.226	-.093	-0.019	20.85	22.26	20.40
每天使用網路時間	-.082	.069	.086	-0.066	4.21	4.59	4.25
接觸網路時間	.019	-.015	.106	.123	5.32	5.16	5.15
每週使用網路時間	.055	.082	-.183	-0.087	29.73	33.28	30.27
學業成績	.024	.079	.052	-0.024	1.78	1.87	1.77

\*\*\* $p < .001$  (+代表有區別能力，++代表有強區別能力)

在第二條區別函數中，敵意的認知曲解的錯誤標籤與目擊網路論戰經驗較能預測「迂迴表達網路敵意群體」，且能區分出「網路論戰沈默群體」。顯示敵意認知曲解的成分的確能促發在網路上敵意之表現，另外目擊其他網路使用者在網路上論戰的經

驗，也會引發無關第三者的敵意。由於此群體是居於第三者的角色，所以在參與論戰時較不會取採衝動或直接的反應，因此傾向以間接的方式表達其不滿。

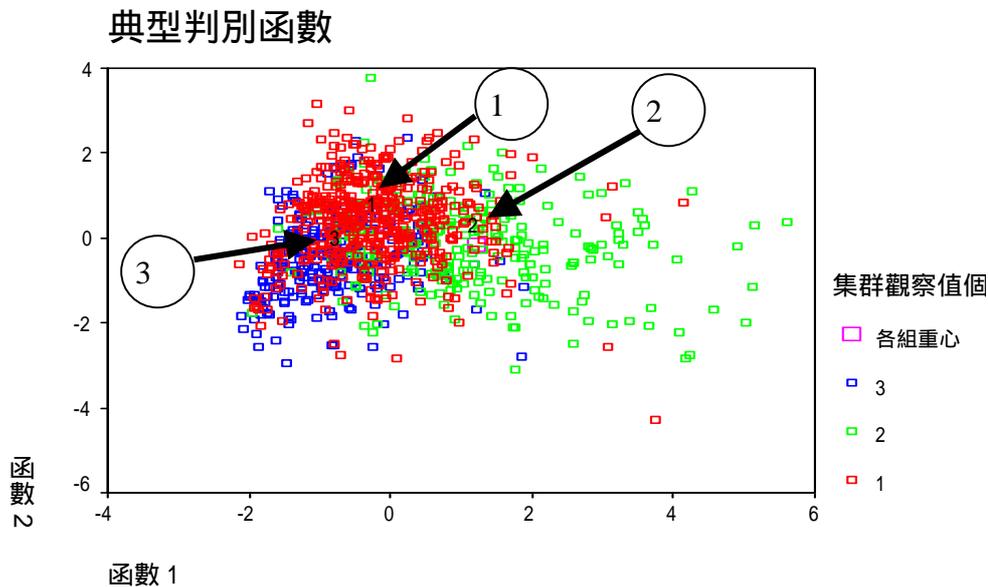


圖1-2 三個網路敵意群體在區別函數之分佈圖

註：1為迂迴表達網路敵意群體，2為直接表達網路敵意群體，3為網路論戰沈默群體。群中心以圓圈標號表示

若以 Fisher's 分類函數係數進行分類預測，表 1-10 顯示其整體正確預測率為 63.1%，各群體正確預測率分別為 78%（迂迴表達網路敵意群體），57.6%（直接表達網路敵意群體），42%（網路論戰沈默群體）。而在交叉驗證中，由表 1-11 可以發現，整體正確預測率為 61.5%，各群體正確預測率分別為 76.3%（迂迴表達網路敵意群體），56.1%（直接表達網路敵意群體），40.5%（網路論戰沈默群體）。

表 1-10 區別分析函數對網路敵意群體之預測結果摘要表

群體別	實際人數	預測精確人數	正確預測率(%)
迂迴表達網路敵意群體	468	365	78%
直接表達網路敵意群體	255	147	57.6%
網路論戰沈默群體	264	111	42%
分組正確預測力 63.1%			

表 1-11 網路敵意群體回推區別分析預測歸類之交叉預測結果摘要表

群體別	實際人數	預測精確人數	正確預測率(%)
迂迴表達網路敵意群體	468	357	76.3%
直接表達網路敵意群體	255	143	56.1%
網路論戰沈默群體	264	107	40.5%
分組正確預測力 61.5%			

## 八、結論與建議

網路上的論戰是否具有獨特性、個殊性，或者只是網路文化使然，導致任何人一進入網路的場域便會忽略社會規範的影響，有許多學者提出不同的看法(Joinson, 1998; Lea & Spears, 1991; Kayany, 1998)。本研究是針對網路論戰之心理層面進行分析，在無法建立實驗情境之困境下，將現實世界中敵意的認知行為模式應用在網路環境中。研究結果發現，網路使用者的網路敵意表現方式並不相同，樣本人群中可區分出三個表達敵意程度不同的群體：「直接表達網路敵意群體」、「迂迴表達網路敵意群體」與「網路論戰沈默群體」，直接表達網路敵意與沈默者比例較低，各約佔四分之一強的比例，約半數的網路使用者傾向於使用間接、迂迴的方式表達網路上不滿的情緒。

過去對網路論戰的研究有的主張網路論戰是普遍的網路文字現象，有的則發現論戰只佔網路極少數比例的文字(Thompson & Foulter, 1996)，這些分析都限於媒體上傳達的語意，如果把參與論戰者的內在心理歷程納入考量，則本研究顯示有四分之一強的網路使用大學生群體屬於直接表達網路敵意的群體，他們對網路人群關係有較高的懷疑、勇於表達憤怒、敢於採用直接口語攻擊、團體攻擊、甚至嘗試其他網路攻擊的行為(如製造垃圾信件、癱瘓對手的信箱等)。因此網路論戰並不是少數文字現象而已，而是有一定的參與人群，未來研究值得循本研究主題繼續追蹤分析。

就網路敵意的認知行為模式而言，遠端因素如：網路上的負面情感與憂鬱，會透過認知曲解的中介，進而影響網路敵意的展現。不過網路易得性與人際焦慮無法如假設模式中所建議的，此二因素對網路敵意不具預測力，可能與操作化定義之不適當有關。不過整體而言，以有限的調查研究資料，已供網路敵意的認知行為模式一個堅強的立論基礎。

由於研究屬於調查研究，尚有許多不足之處有待克服。首先，在推論性方面，本研究只選取北部三所大學，雖然目前台灣大專院校學術網路之建置已無地域上之差異，但是並無實證研究指出不同地區的大專學生在網路使用上之差異，因此在進行全國性的推論時須謹慎小心。其次，本研究屬於橫斷性的問卷調查，只能針對目前的現況加以描述，無法深入探討網路敵意的深層狀態，建議往後研究可以針對個案進行質化訪談，或者收集網路上的論戰文字敘述進行內容分析，在交叉檢驗下更能增加本研究的效度。最後，受限於時間與研究者之能力，無法發展一套有效的輔導方案並檢驗其成效，不過在探索性的研究方面已有初步的結果，提供日後的研究發展的新方向。因此針對網路去抑制的行為表現，建議學校能設計相關的課程，教導學生正確的網路使用規範，並且重建學生使用網路時不合理之認知，使其擁有一個健全、無威脅的網路環境，以及培養正確的網路使用態度。

## 第二部分：自動化網路敵意文章分級系統之初探研究

### 一、前言

由於網路的普及與方便性，許多的個人及團體紛紛將文件放置在網路上，提供需要者快速的取用途徑，當使用者要找尋某一類主題文件時，便可以利用搜尋引擎，快速找到所需要的資料。但由於網路上的文件數量極多，因此在尋找所要資訊時所遭遇到的問題，也從原來的不易取得文件，變成不易找到所需要文件(Belkin,1992)。為了讓使用者可以很快速的找出想要的文件，我們通常會先將文件分類放置(Croft & Larkey, 1996; Liu & Yang, 1999; Tsay & Wang, 2000; Yang, 2001; Chao & Wai, 1998)，當使用者有查詢文件的需要時，便可以利用搜尋引擎，在搜尋引擎中輸入與文件相關的關鍵字或關鍵句，但是這樣的查詢方式主要是針對一些類別較明顯的文件，例如：要查詢關於電腦硬體的文件，可以直接輸入電腦硬體、主機板、或電腦週邊等關鍵字，即可找到相關文件。

但如果想要查詢的文件是屬於類別較不清楚的，例如想要找出自傳，或者是表達謝意的文件，相關的討論就不是那麼多了。在公開的論壇或是留言版上，常常會看到一些惡意攻擊或是謾罵的文件，這類文件通常會偏離主題甚多，且容易引起網路上的論戰(flamming)，而網管人員也常常要花費許多時間處理這些文件，因此若能設計一個能自動分類敵意文章的系統，就可幫助網管人員快速篩選出具有敵意的文章，減少網管人員處理這類文章的時間，提昇管理效率。本文的目的，即是希望能設計一個自動化敵意文章分級系統，由使用者先提供一些具有敵意的文章（以下簡稱訓練文件），系統會依照所提供的文章，找出想要搜尋文章類型的特徵，並利用此特徵計算文件庫中所有文件的特徵值，計算出每篇文件與敵意文章的相似度，並設定門檻，最後將敵意文章分類出來。

### 二、自動化敵意文章分級系統之建置步驟

本研究的主要工具為一套建立在 web 上的敵意文章分類系統，系統環境如表 2-1 所示：

表 2-1：作業環境設置

作業系統	RedHat Linux 7.2 (Enigma)Kernel 2.4.7-10 on an i686
web 伺服器	Apache 1.3.20-16
資料庫	Postgresql 7.1.3-2
程式語言	PHP-4.0.6-7、PHP-pgsql-4.0.6-7
系統網址	<a href="http://163.25.180.120/cgi/nctu/i_r/index.html">http://163.25.180.120/cgi/nctu/i_r/index.html</a>

為了在實驗中觀察各步驟的影響以及控制變因，將敵意文章分類系統切割為數個

步驟實施，系統細項功能如下表所示：

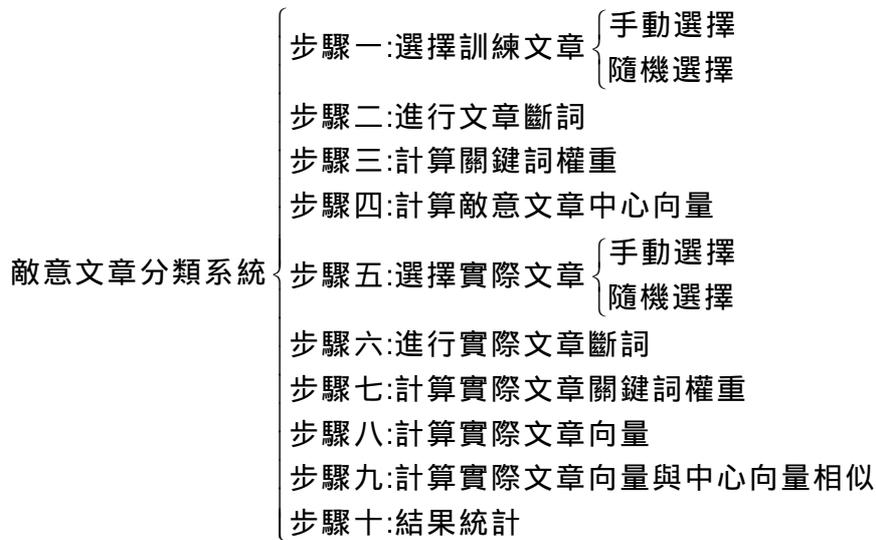


圖 2-1：敵意文章分類系統實行步驟

步驟一與步驟二的主要目的是要建立能代表文章的關鍵詞詞庫，建立方式則是從訓練文件中找出文件特徵(即關鍵詞)，放入關鍵詞詞庫中，常用來建立詞庫的方法有兩種，辭典式斷詞法與統計式斷詞法，辭典式斷詞法需事先將能判別文件類別的語詞放入詞庫中，如中央研究院詞庫小組(<http://godel.iis.sinica.edu.tw/CKIP/>)所建立的語料庫，但由於有些類別文件在定義上並不是那麼的清楚，以這次主題為例，要找出具敵意文件的關鍵詞並不容易，況且由於地區、時間、討論主題的不同，即使他們都是具有敵意的文件，慣用語也會有不同的地方，而統計式斷詞法則是先蒐集一些同類型的文件(以下稱為訓練文件)，經過斷詞後，計算出語詞的權重，並將可以代表此類文件的語詞選為關鍵詞，並放入詞庫中，此方法可以解決辭典式斷詞法的缺點，因此在本系統中將採用統計式斷詞法。

而取出語詞的長度，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〔22〕，因此取出語詞的長度將以長度為 2 的二連字詞為主。在電腦系統中，由於大小寫英文字母再加上常用符號，沒有超過 128 個，因此是以一個 byte 來儲存字母，但是常用中文字就將近 5000 個，所以必須用 2 個 byte 來儲存中文，但是這樣的方式，在斷詞時，會造成當大的困難，以要斷詞的長度為兩個字為例，如果一份文件中，只有英文字母和符號，在斷詞時的處理，只要從文件一開始，每次擷取 2 個 byte 的資料到文件結束即可將整份文件斷詞完成，但如果一份文件中含有中文及英文，則必須從文件一開始，先判別第一個 byte 的二進位碼是否大於 128，如果不是，再判別第二個 byte 是否大於 128，如果也不是，表示取出的資料不包含中文，故只要直接 2 個 byte 資料即可，但如果的是的話，表示取出的資料，第一個字為英文或符號，第二個字為中文，因此總共需取出 3 個 byte 的資料，但是如果第一個 byte 大於 128，則需判別第三個 byte 的二進位碼是否大於 128，如果不是，表示取出的資料，第一個字為中文，第

二個字為英文或符號，如果是的話，表示這 4 個 byte 總共包含 2 個中文字，因此需取出 4 個 byte 的資料，整個程序如圖 2-2 所示：

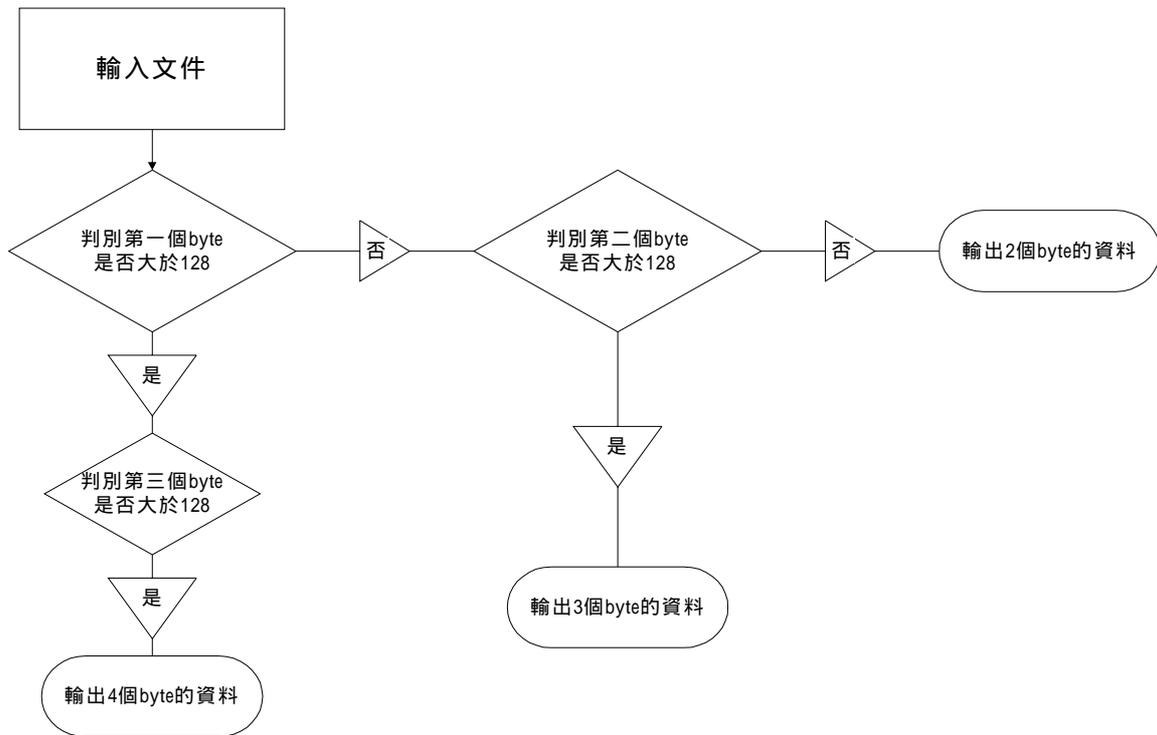


圖 2-2：中文關鍵詞斷詞流程(未導入 iconv 函數)

中文的斷詞比英文的斷詞要來的複雜許多，若要處理大量文件的斷詞，會耗費大量時間，由於本研究主要針對中文文件，因此本文中的系統將只擷取文章中的中文字，並利用 iconv 函數(<http://www.iconv.com>)，來加快斷詞的速度。擷取流程如圖 2-3：

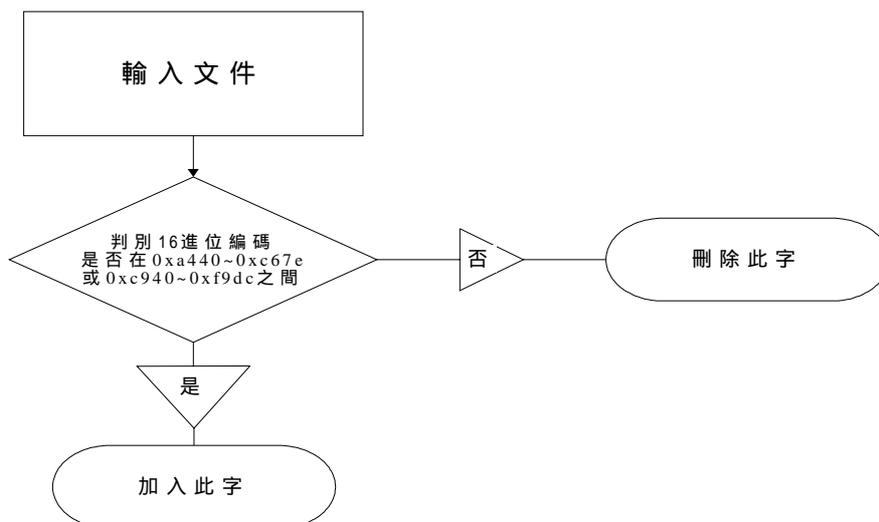


圖 2-3：中文關鍵詞斷詞流程(導入 iconv 函數)

以下列文章為例：

表 2-2：未經斷詞處理前的文章內容

作者	夏天
群組	tw.bbs.comp.hardware
標題	Re: 微星主機板作弊被抓包了..
時間	2003-06-14 05:04:16
<p>微星主機板作弊被抓包了..</p> <p>引述《(甲蟲)》之銘言：</p> <p>&gt; 你們為何不買 ASUS P4C800 啊?</p> <p>&gt; .</p> <p>    P4C800 除了賣那顆 875 還有什麼?</p> <p>南橋不搭 ICH5R 要用 20378 作 S-ATA raid</p> <p>那想用 IDE raid 的還要另買轉接頭?</p> <p>NIC 也是，明明有 CSA 卻用 3Com</p> <p>看了就不爽</p> <p>除了 P4P800-D 看起來比較不錯之外(但是 VIA 的 raid...-_-?)</p> <p>ASUS 其他 865/875 的板子都沒興趣</p> <p>--</p> <p>Origin: 精靈之城      From: vai.dorm4.ntnu.edu.tw</p>	

資料來源：tw.bbs.comp.hardware

系統會先將文章中的非中文文字，也就是標點符號、單位、英文字母、數字移除，移除後的文章內容如下表：

表 2-3：經斷詞處理後的文章內容

作者	熱@@
群組	tw.bbs.comp.hardware
標題	Re: 微星主機板作弊被抓包了..
時間	2003-06-14 05:04:16
<p>微星主機板作弊被抓包了引述風之銘言你們為何不買啊除了賣那顆還有什麼南橋不搭要用作那想用的還要另買轉接頭也是，明明有卻用看了就不爽除了看起來比較不錯之外但是的其他的板子都沒興趣精靈之城</p>	

資料來源：tw.bbs.comp.hardware

文章中除了使用者發表的本文外，通常還包含關於此主題的引言、及簽名，由於這兩部分無明顯的分界符號，或是較一致的特徵，因此並不容易用電腦進行精確的移除工作，所以在本系統中，對引言及簽名檔不進行移除的動作。在非中文文字移除後，系統會對此篇文章進行雙連字詞的擷取動作，並統計相關數據。雙連字詞擷取的結果如下表所示：

表 2-4：文章經斷詞後取出之關鍵詞列表

微星、星主、主機、機板、板作、作弊、弊被、被抓、抓包、包了、了引、引述、述風、風之、之銘、銘言、言你、你們、們為、為何、何不、不買、買啊、啊除、除了、了賣、賣那、那顆、顆還、還有、有什、什麼、麼南、南橋、橋不、不搭、搭要、要用、用作、作那、那想、想用、用的、的還、還要、要另、另買、買轉、轉接、接頭、頭也、也是、是明、明明、明有、有卻、卻用、用看、看了、了就、就不、不爽、爽除、除了、了看、看起、起來、來比、比較、較不、不錯、錯之、之外、外但、但是、是的、的其、其他、他的、的板、板子、子都 都沒、沒興、興趣、趣精、精靈、靈之、之城
--

步驟三、計算關鍵詞權重：在步驟二取出關鍵詞後，接下來要計算詞庫中每個關鍵詞的權重，在本系統中，關鍵字權重是以 TF-IDF 方式取得，先假設關鍵字有  $m$  個， $K$  為關鍵詞集合，

$$K = \{k_1, k_2, k_3, \dots, k_m\} \quad (1)$$

而訓練文件有  $n$  份，將訓練文件集合設為  $D$ ，則

$$D = \{\bar{d}_1, \bar{d}_2, \bar{d}_3, \dots, \bar{d}_n \mid \bar{d}_i \in \bar{V}_m, \forall i = 1, 2, \dots, n\} \quad (2)$$

在訓練文件中出現關鍵字  $k_j$  的篇數為  $n_j$ ，設為

$$N = \{n_1, n_2, n_3, \dots, n_m\} \quad (3)$$

$$\begin{aligned}
 w_{ij} &= tf_{ij} \cdot idf_{ij} \\
 &= \frac{freq_{ij}}{\max_l freq_{il}} \cdot \log \frac{n}{n_j}
 \end{aligned} \quad (4)$$

在式子(4)中， $freq_{ij}$  表示關鍵字  $k_j$  在訓練文件  $d_i$  中出現的次數， $freq_{ij}$  為正整數。 $\max_l freq_{il}$  表示在訓練文件  $d_i$  中出現最多次的關鍵字的次數， $\max_l freq_{il}$  為正整數。若關鍵詞  $k_j$  未在某一文件  $d_i$  中出現，則  $freq_{ij} = 0$ ，因此  $w_{ij} = 0$ 。

(2)、計算出每份訓練文件的向量：

在向量模型(vector model)中，每份文件皆以一個 m 維度的向量，

$$\vec{d} = (w_1, w_2, w_3, \dots, w_m) \quad (5)$$

來表示 [16]，其中  $w_i$  表示關鍵字  $k_i$  在文件  $d$  中的權值重(term weight)，每一份訓練文件可以一個 m 維度的向量來表示，

$$\begin{aligned} \vec{d}_1 &= (w_{11}, w_{12}, w_{13}, \dots, w_{1m}) \\ \vec{d}_2 &= (w_{21}, w_{22}, w_{23}, \dots, w_{2m}) \\ &\quad \bullet \\ &\quad \bullet \\ &\quad \bullet \\ \vec{d}_n &= (w_{n1}, w_{n2}, w_{n3}, \dots, w_{nm}) \end{aligned} \quad (6)$$

其中  $w_{ij}$  表示關鍵字  $k_j$  在文件  $d_i$  中的權值重(term weight)，

$$W = \{w_{ij} \mid i = 1..n, j = 1..m\} \quad (7)$$

步驟四、計算敵意文章中心向量：將所有訓練文件的向量平均後即得到敵意中心向量(hostility center vector)，以下簡稱 hcv。

設  $\vec{C}$  為 hcv，則

$$\begin{aligned} \vec{C} &= (c_1, c_2, c_3, \dots, c_m) \\ &= \left( \frac{\sum_{i=1}^n w_{i1}}{n}, \frac{\sum_{i=1}^n w_{i2}}{n}, \frac{\sum_{i=1}^n w_{i3}}{n}, \dots, \frac{\sum_{i=1}^n w_{im}}{n} \right) \end{aligned} \quad (8)$$

步驟五、選擇實際文章：此步驟可以使用者需求，選擇自己所需要的實際文章，或指定文章類別與篇數兩個參數，由系統隨機選取文章：

步驟六到步驟八、計算實際文件向量：假設現在有一份文件  $r$ ，則此文件在文件向量空間中會以一個 m 維度的向量來表示

$$\vec{r} = (r_1, r_2, r_3, \dots, r_m) \quad (9)$$

$r_i$  表示關鍵字  $k_i$  在文件  $r$  中的權重、權重的算法，則以 Salton and Buckley 所建議的計算方式

$$r_i = (0.5 + \frac{0.5 \text{freq}_{ir}}{\max_l \text{freq}_{lr}}) \log \frac{n}{n_i} \quad (10)$$

$\forall i = 1 \dots m$

其中  $\text{freq}_{ir}$  表示關鍵字  $k_j$  在實際文件 r 中出現的次數， $\max_r \text{freq}_{ir}$  表示在實際文件 r 中出現最多次的關鍵字的次數，而 n 與  $n_i$  則與之前的定義相同。

步驟九、計算實際文章向量與敵意中心文章向量相似度：實際文章與敵意中心文章向量的相似度，則需利用餘弦函數來計算，文件 r 與敵意文章的相似度以  $\text{sim}(\bar{r}, \bar{c})$  來表示，函數 sim 定義如下：

$$\text{sim}(\bar{r}, \bar{c}) = \frac{\bar{r} \cdot \bar{c}}{|\bar{r}| |\bar{c}|} = \frac{\sum_{i=1}^m r_i c_i}{\sqrt{\sum_{i=1}^m r_i^2} \sqrt{\sum_{i=1}^m c_i^2}} \quad (11)$$

sim 函數介於 0 與 1 之間，數值越大，表示實際文章與 hcv 的相似度越高。

### 三、敵意文章分級系統之初階實驗

本實驗研究步驟分為四個階段、第一階段為實驗前之取樣及樣本分析：

#### (一) 研究樣本選取：

本實驗的訓練文件取樣來源為實際 BBS 文章中被分類為論戰文章之文章，包含分歧(21 篇)、爭論 (143 篇)、緊張(145 篇)、敵對(115 篇)、尖銳敵對(6 篇)，合計共 430 篇文章。

表 2-5：實驗文件取樣來源

取樣來源	台灣 BBS 站的硬體版 tw.bbs.comp.hardware
取樣方式	本實驗採連續抽樣，所有樣本的討論版文章之上傳時間為 2003-06-06 04:59:05 至 2003-06-18 15:38:48。
文章長度	樣本文章長度介於 0 與 1371 之間。為了避免文章長度影響文章自動化敵意值，因此本文所用分類方法將針對文章長度作正規化。
備註	本實驗只判斷文章是否具有敵意，因此若有不符合本版主題的文章，如廣告或發錯版面文意，皆不予刪除。

#### (二) 實驗前之前置設定：

對訓練文章與實際文章進行敵意相關程度的設定，本實驗以一位具有網路學習碩

士學位之國中教師(擔任資訊組長已多年)為判斷者 A，另外兩位為判斷者 B 與 C 皆為現職國中教師且擔任資訊組長職務均超過 3 年，由於資訊組長需要擔任學校網站與討論板管理的職務，對判定敵意文件都具有相當多的經驗，因此可提高相關認定的一致性與準確度。三位判斷者先各自片段敵意層級，意見不同時，由判斷者 A 主動邀請協商，以協商結果為該文章的敵意等級。

具敵意的文章是由論戰中產生而來，但研究者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高者，則較容易觀察，因此本文中對具敵意文章的認定，主要是先利用文章篇數出現較多的特性，找出論戰文章，再從論戰文章中，依照 Thompsen and Foulger 提出的論戰五種過程。將文章的敵意設為 6~10，數字越高，代表越具有敵意。

在敵意程度的評定方式，本實驗採取二階段評定，第一階段先由三位判斷者給予每篇文章 1 分或 0 分，1 分表示判斷者判定此篇文章具有敵意，0 分則否，三位判斷者給定的分數設為  $S_A$ 、 $S_B$ 、 $S_C$ ， $S$  則為三個分數的總和， $S = S_A + S_B + S_C$ ，由於在多位判斷者判定文章類別時，對於非主題的判定一致性較高，因此先將  $S=0$  的文件判定為非敵意文章，再由判斷者 A 將非敵意文章分為不完整文章、廣告文章、其他版面文章、發問或回覆文章，而  $S>0$  的文章則判定為具有敵意文章，由判斷者 A 將敵意文章作進一步的確認，類別則以 Thompsen and Foulger 的分類為基礎，分為分歧、爭論、緊張、敵對、尖銳敵對五類，敵意判定流程如圖 2-4。

先將 5000 篇文章以人工方式給定敵意值，給定方式為先將所有文件分類，先粗分有敵意文章、無敵意文章、及其他類文章：

#### 1. 其他文章：

- (1) 不完整文章：此類文章的特性為文章長度太短，無法呈現某個完整的概念，或是由於其他錯誤，導致文章不完整，例如撰寫文章時誤按送出，或是網路連線中斷導致不完整的文章。
- (2) 廣告文章：文章內容為宣傳某樣產品或有交易行為之文章。
- (3) 其他版面文章：本次實驗所使用文章為 tw.bbs.comp.hardware 的文章，主要內容為討論電腦硬體的文章，若文章主題不符，如討論文學或政治，則歸類為其他版面文章。

2. 敵意文章：由前面的討論，我們發現敵意文章是在論戰過程中產生的，當某些因素引起使用者敵意後，使用者會將其敵意藉由文章表現出來，但是並無法得知敵意是在論戰的哪一個時期產生，而論戰文章由於篇數較多，較容易被觀察，因此先計算對於某一主題的文章篇數，若文章篇數明顯高於討論其他主題的文章篇數，則表示此一主題文章屬於論戰文章，並將其初步定為具有敵意的文章。

3. 非敵意文章：此類文章包含一般討論文章，即發問與討論。

分類方式如表 2-6 所示，經由此分類方式分類後，各文章的篇數如表 2-7。

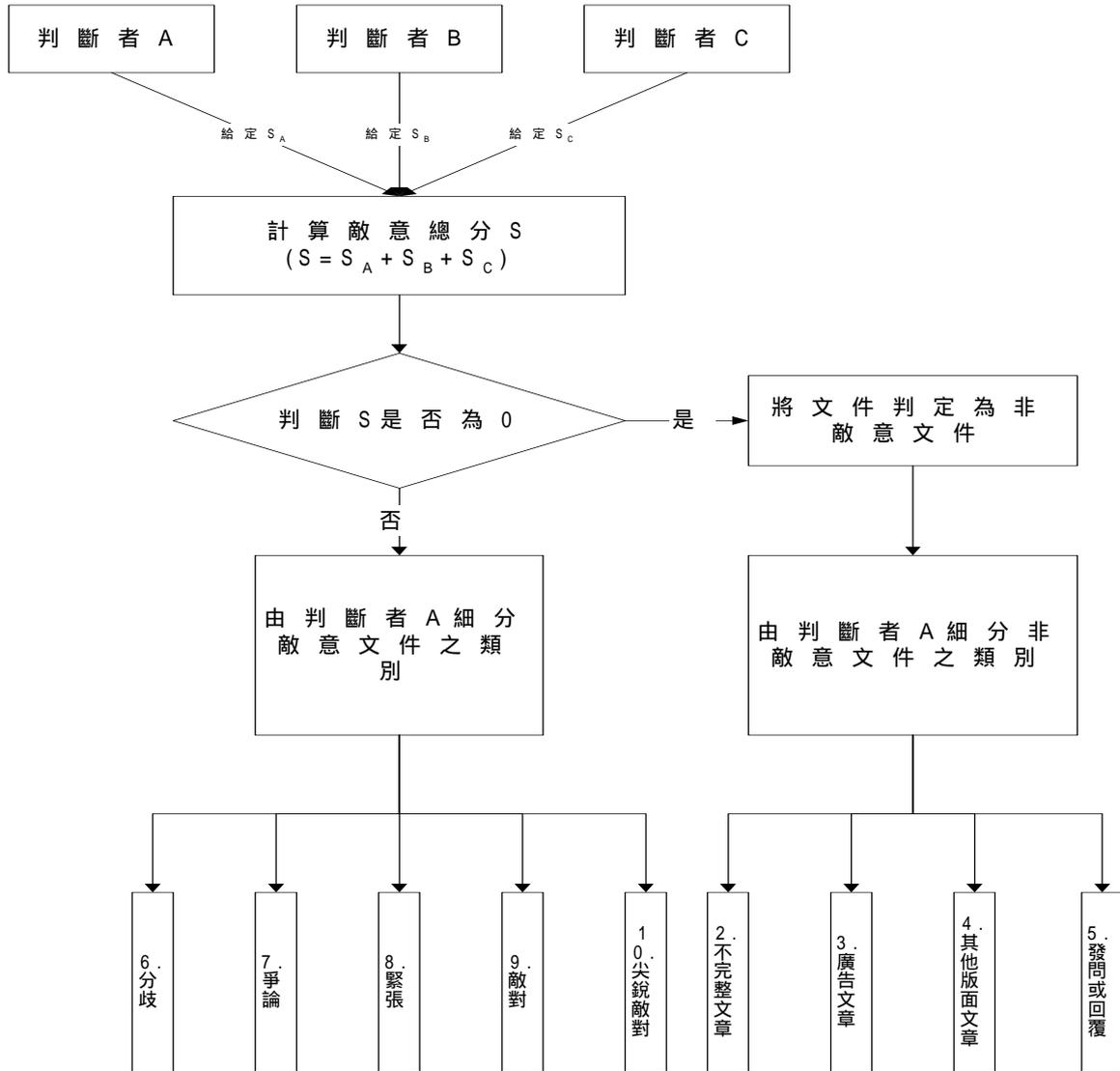


圖 2-4：文章敵意判別流程

表 2-6：所有文章分類表

5000篇tw.bbs.comp.hardware的文章

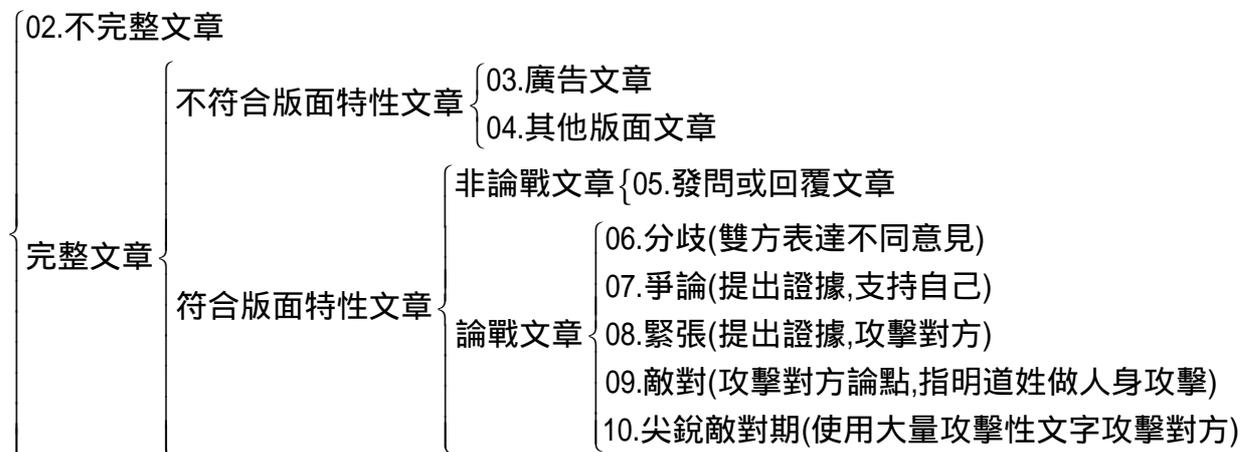


表 2-7：各類文章篇數及所佔比例

類別代碼	文章類別	篇數	所佔比例
2	不完整文章	45	0.90
3	廣告文章	78	1.56
4	其他版面文章	5	0.10
5	發問或回覆文章	4442	88.84
6	分歧	21	0.42
7	爭論	143	2.86
8	緊張	145	2.90
9	敵對	115	2.30
10	尖銳敵對	6	0.12

而文章類別代碼為 6、7、8、9、10 的文章屬於論戰文章，共有 12 個討論主題，其中關於各主題的文章類別分布如表 2-8：

表 2-8：論戰文章中各主題所佔篇數及比例

序號	主題代碼	討論主題	類別代碼	篇數	比例
1	A	[問題]HITACHI 的硬碟	7	2	8.60%
			8	5	
			9	25	
			10	5	
2	B	Lite-on 真好	8	1	0.23%
3	C	台中市哪裡賣電腦的便宜又好的??	8	1	0.70%
			9	2	
4	D	台灣...好貴的"寬頻" >"<	6	3	48.14%
			7	83	
			8	73	
			9	48	
5	E	全世界只有台灣收什麼鬼電路費	8	2	8.60%
			9	18	
			10	17	
6	F	全國產電腦，可能嗎?	6	6	10.93%
			7	30	
			8	9	
			9	2	
7	G	如何讓 3DMark2001 的分數破萬??	8	1	1.40%
			9	4	
			10	1	
8	H	青雲的產品網頁沒有中文	6	3	8.60%

			7	20	
			8	15	
9	I	現在的顯示卡	8	1	0.23%
10	J	微星主機板作弊被抓包了..	6	12	10.93%
			7	5	
			8	20	
			9	10	
11	K	請問 so-net 的連線品質好嗎	7	5	1.16%
12	L	請問一下 k6-2 的 cpu	9	1	0.23%
總計				430	100%

### (三) 門檻值設定：

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，數值越大表示與敵意文章越相似，反之則否，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，而門檻值越高，判定的精確度越高，但有越多的文章會被判定為非敵意文章，反之，若門檻值越低，則能找出更多敵意文章，但也會因此而降低精確度，因此必須先進行門檻值實驗，找出較佳的敵意門檻值，同時提高判定敵意文章與非敵意文章的精確度。

為了解以不同主題為訓練文章時的相似度分布，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，此六個主題為 A( [問題]HITACHI 的硬碟)、D (台灣....好貴的"寬頻" >"<)、E(全世界只有台灣收什麼鬼電路費)、F(全國產電腦，可能嗎?)、H(青雲的產品網頁沒有中文)、J(微星主機板作弊被抓包了..)，我們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 10、20、30 篇文章，進行敵意相似度的計算，詳細流程如下圖所示：

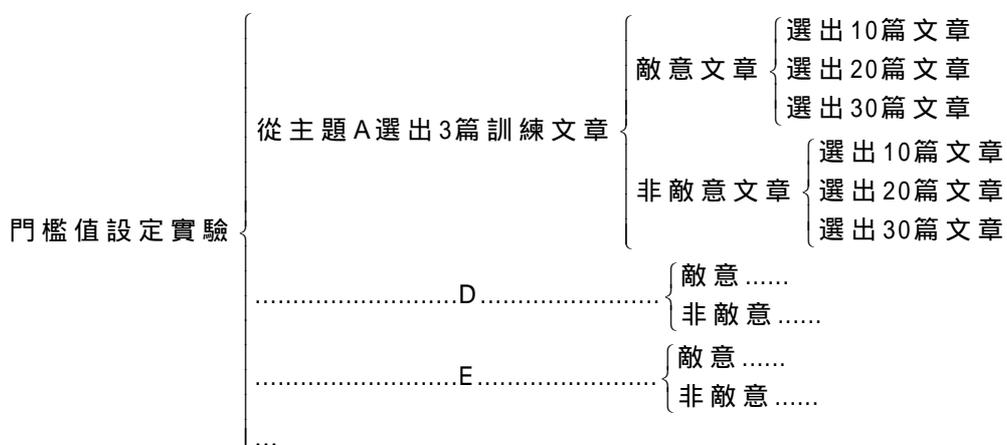


圖 2-5：門檻值實驗流程

實驗後的各主題平均值如下表所示：

表 2-9：各組平均值

	主題 A	主題 D	主題 E	主題 F	主題 H	主題 J
敵意 10	0.210303	0.160904	0.241688	0.108914	0.273305	0.147637
非敵意 10	0.179063	0.099368	0.115494	0.090258	0.123775	0.106237
敵意 20	0.223465	0.161734	0.290873	0.132015	0.208256	0.139688
非敵意 20	0.147783	0.085326	0.118047	0.098903	0.128751	0.099666
敵意 30	0.194430	0.164296	0.244877	0.116721	0.216210	0.138887
非敵意 30	0.139652	0.092839	0.120687	0.096279	0.131310	0.107615

### 1. 同一主題時，敵意與非敵意文章相似度差異檢定

敵意文章的門檻值分布在 0.290873 與 0.108914 之間，而非敵意文章的門檻值分布在 0.179063 與 0.085326 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，為了了解在同一主題時，計算出的敵意文章與非敵意文章的相似度，是否具有明顯差異，以及因此先假設  $H_0$  為 ” 訓練文章為同一主題時，所計算出的敵意文章與非敵意文章的相似度，沒有明顯差異 ”， $H_1$  為 ” 訓練文章為同一主題時，所計算出的敵意文章與非敵意文章的相似度，具有明顯差異 ”，並利用獨立樣本 t 檢定來進行檢定，結果如表 2-10，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 2-10：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
10 篇	0.230	0.011	0.001	0.107	0.012	0.004
20 篇	0.000	0.000	0.000	0.001	0.000	0.004
30 篇	0.002	0.000	0.000	0.048	0.000	0.004

### 2. 不同主題時，計算出的敵意文章相似度差異檢定

以不同的主題作為訓練文章，來計算敵意文章的相似度，是否也會產生差異？我們的虛無假設為不同的主題作為訓練文章，來計算敵意文章的相似度沒有差異。利用單因子變異數分析(one-way ANOVA)來進行假設檢定，結果如表 15 所示：

表 2-11：以主題別作為因子之單因子變異數分析摘要表

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		

\*\*\* $p < .001$

差異檢定達顯著性，拒絕虛無假設，即以不同的主題作為訓練文章時，所計算出之敵意文章的相似度具有明顯差異，六個主題的敵意平均值折線圖如圖 2-6 所示：

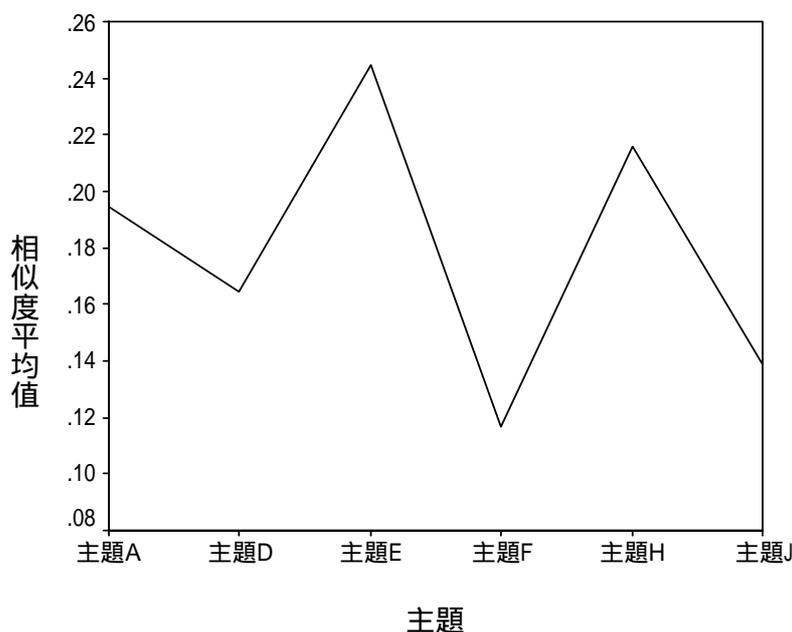


圖2-6：主題與非敵意文章敵意平均值關係折線圖

### 3. 第三階段 (實驗後結果分析)：

為了解此方式所計算出之文件相似度與人類思考方式之差異，採取實驗法，以系統判讀為實驗組，人工判讀為對照組，將電腦所判別出來的相似度與人工判讀的相似度做對照。精確率(precision)及召回率(recall)是常用來評估文件分類系統效能的兩個數值，精確率(precision)的意思是，所有擷取出的文件中，與搜尋 q 有關的文件比率。而召回率(recall)的意思是，所有與搜尋 q 有關的文件，能被擷取出的比率。例如現在共有 100 篇文件，其中與搜尋 q 有關的文件共有 40 篇，假設某次的擷取共取出 50 篇，而此 50 篇文件中有 30 篇與搜尋 q 有關，則：

$$\left\{ \begin{array}{l} \text{精確率}(precision) = \frac{30}{50} = 0.6 \\ \text{召回率}(recall) = \frac{30}{40} = 0.75 \end{array} \right. \quad (12)$$

由於實驗設計中文章分類只有敵意與非敵意兩種，必須對評估指標作修正，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。

$$\left\{ \begin{array}{l} HR = \frac{\text{人工與系統皆認為具有敵意的文章篇數}}{\text{抽出文章總篇數}} \\ NHR = \frac{\text{人工與系統皆認為不具有敵意的文章篇數}}{\text{抽出文章總篇數}} \end{array} \right. \quad (13)$$

以上述實驗方式為例，假設在某次實驗中取出 n 篇文章，經由實驗後，人工與系統對認定文章是否有敵意的組合，可能有下列四種：

- (1) 人工敵意 = 1，系統敵意 = 1，人工與系統皆認定此篇文章具有敵意。
- (2) 人工敵意 = 1，系統敵意 = 0，人工認為具有敵意，系統不認為具有敵意。
- (3) 人工敵意 = 0，系統敵意 = 0，人工認為不具敵意，但電腦認為具有敵意。
- (4) 人工敵意 = 0，系統敵意 = 0，人工與系統皆認為不具有敵意。

$$\text{則} \left\{ \begin{array}{l} HR = \frac{a}{n} \\ NHR = \frac{d}{n} \end{array} \right. \quad (14)$$

表 2-12：人工與系統對敵意認定的可能組合

組合代碼	人工	系統	篇數
1	1	1	A
2	1	0	B
3	0	1	C
4	0	0	D
			N

#### 4. 實驗結果：

非敵意文章的相似度分布在 0 到 0.1695 之間，平均為 0.109569，各敘述統計量如表 2-13 所示：

表 2-13：非敵意文章，每次取樣 10 篇，進行 10 次實驗之敘述統計量

最小值	最大值	平均數	標準差
0	.17	.11	.036

由於門檻值為0.17，因此所有文章皆被判別為非敵意文章，NHR值皆為1，與人工認定方式完全符合。各次實驗的NHR值如表2-14所示：

表 2-14：非敵意文章，每次取樣 10 篇，進行 10 次實驗之實驗結果

id	a	b	c	d	NHR
1	0	0	0	10	1
2	0	0	0	10	1
3	0	0	0	10	1
4	0	0	0	10	1
5	0	0	0	10	1
6	0	0	0	10	1
7	0	0	0	10	1
8	0	0	0	10	1
9	0	0	0	10	1
10	0	0	0	10	1

敵意文章的相似度分布在 0.091 到 0.2135 之間，平均為 0.152683，各敘述統計量如表 2-15 所示：

表 2-15：敵意文章，每次取樣 10 篇，進行 10 次實驗之敘述統計量

最小值	最大值	平均數	標準差
.091	.214	.153	.024

只有少部分文章皆被判別為敵意文章，NR 介於 0.1 與 0.4 之間，平均值為 0.25。各次實驗的 HR 值如表 2-16 所示：

表 2-16：敵意文章，每次取樣 10 篇，進行 10 次實驗之實驗結果

id	a	b	c	d	HR
1	3	0	7	0	0.3
2	1	0	9	0	0.1
3	2	0	8	0	0.2
4	3	0	7	0	0.3
5	3	0	7	0	0.3
6	4	0	6	0	0.4

7	3	0	7	0	0.3
8	1	0	9	0	0.1
9	1	0	9	0	0.1
10	4	0	6	0	0.4

非敵意文章取樣 20 篇時，敵意文章的相似度分布在 0 到 0.213 之間，平均值為 0.108197，各敘述統計量如表 2-17 所示：

表 2-17：非敵意文章，每次取樣 20 篇，進行 10 次實驗之敘述統計量

最小值	最大值	平均數	標準差
0	.213	.108	.037

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。各次實驗的 NHR 值如表 2-18 所示：

表 2-18：非敵意文章，每次取樣 20 篇，進行 10 次實驗之實驗結果

id	a	b	c	d	NHR
51	0	1	0	19	0.95
53	0	0	0	10	1
58	0	0	0	20	1
59	0	1	0	19	0.95
60	0	0	0	20	1
61	0	0	0	20	1
62	0	2	0	18	0.9
63	0	1	0	19	0.95
64	0	0	0	20	1
65	0	0	0	20	1

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 0.155687，各敘述統計量如表 2-19 所示：

表 2-19：敵意文章，每次取樣 20 篇，進行 10 次實驗之敘述統計量

最小值	最大值	平均數	標準差
.047	.312	.156	.035

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。各次實驗的 HR 值如表 2-19 所示：

表 2-19：敵意文章，每次取樣 20 篇，進行 10 次實驗之實驗結果

id	a	b	c	d	HR
1	9	0	11	0	0.45
2	9	0	11	0	0.45
3	6	0	14	0	0.3
4	5	0	15	0	0.25
5	2	0	18	0	0.1
6	6	0	14	0	0.3
7	3	0	17	0	0.15
8	6	0	14	0	0.3
9	6	0	14	0	0.3
10	6	0	14	0	0.3

在分類非敵意文章時，每次取樣 10 篇，進行 10 次實驗之 NHR 的值皆為 1，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。因此我們以門檻值為自變項，HR 與 NHR 的平均值為依變項，來觀察準確率的變化情形。

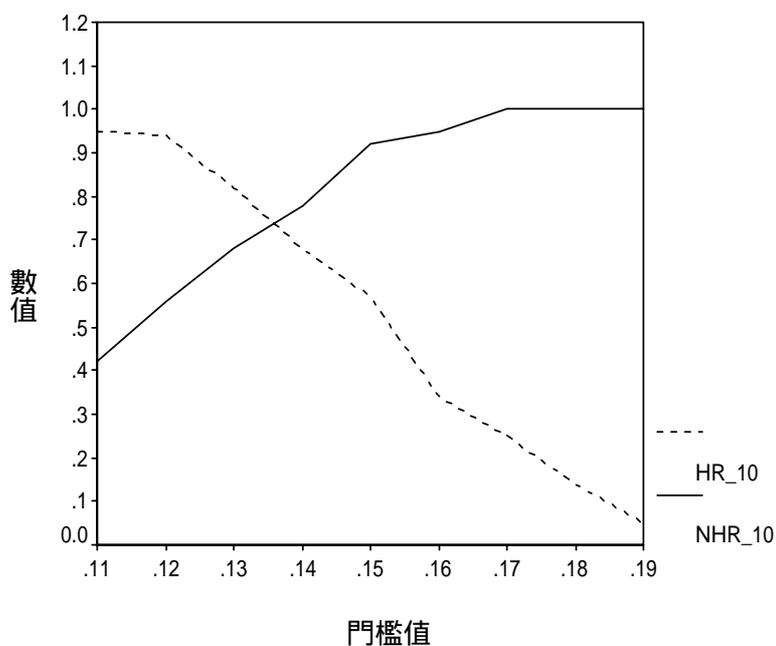


圖 2-7：取樣 10 篇時，門檻值與 HR、NHR 之關係折線圖

#### 四、結論

本研究進行敵意文章分級系統的初探，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. 同一主題時，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，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. 不同主題時，計算出的敵意文章相似度，具有顯著差異。
3. 不同主題時，計算出的非敵意文章相似度，具有顯著差異。
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。

本研究有下列研究限制：

1. 研究設計的限制：本系統設計主要是針對中文文章進行分類，如權重計算方式或是斷詞方式亦針對中文的特性設計，無法確定對於其他語文文章，亦具有相同效果。
2. 研究對象限制：本研究因時間、人力的限制，只採用 tw.bbs.comp.hardware 版面上的文章，作為研究對象，文章的取樣篇數亦只有 5000 篇，樣本並不多，再由樣本分析來看，取樣方式為按照時間的連續抽樣，因此對實驗結果亦會產生影響，在推論研究結果時，必須考慮這些因素。
3. 研究應用的限制：本實驗對象主要是 tw.bbs.comp.hardware 版面上的文章，文章具有硬體討論版的特性，無法保證亦適用在其他版面。

建議未來研究方向：

1. 降低 HCV 的維度：在計算 hcv 時，隨機抽樣出 10 篇文章進行實驗，所需時間約 2 小時 20 分左右，計算一篇文章相似度的平均時間約 14 分左右，未來研究可朝向如何在不影響分類成效的前提下，降低 hcv 的維度，以提昇分類效率。
2. 敵意文章組成方式：本研究並未對會形成敵意文章的二連字詞進行研究，僅在附錄中列出出現次數高於 10 次的二連字詞，未來可研究敵意文章中某些特別語詞是否出現頻率較高，以深入了解敵意文章組成。
3. 自動取得最佳門檻值：本研究是以半自動的方式取得門檻值，需要先進行門檻值實驗，求出第一次最佳門檻值後，在真實環境中進行實驗，觀察分類結果，並進一步調整較佳門檻值，未來可研究將門檻值設定程序自動化的方法，

降低人為誤差，以提昇分類成效。

4. 敵意文章的定義：本研究為敵意文章分級系統初探，在進行文章分級時，僅透過敵意的外顯行為，也就是論戰，來將敵意與非敵意文章進行初步分類，未來可研究針對敵意文章進行直接的定義或是給定操作型定義，以便針對敵意文章，做進一步的分級。
5. 不同主題所造成 hcv 的差異：本研究僅利用統計方法，證明不同主題所形成的 hcv 具有差異，但並未針對造成差異的原因，進行質化研究，未來可研究造成不同主題 hcv 產生差異的原因，以便找出較能代表敵意文章的 hcv。
6. 心理敵意與敵意文章之關聯：敵意是人類心理的狀態，而敵意文章則是敵意的外顯行為，未來可研究兩者之間的關聯性，也就是個體在具有敵意的狀態下，是如何產生敵意文章，以及敵意文章是如何使個體產生敵意，來釐清敵意與非敵意文章之間的關聯性。
7. 減少敵意文章之機制：大量的敵意文章，常會造成網管人員的困擾，也會提高個體產生衝突的可能性，因此一般的管理模式，僅能被動的找出敵意文章後，將其刪除，但其負面影響通常已產生，未來可研究敵意文章自動消除的機制，在敵意文章出現之初即由系統進行消除的動作，以減少負面影響。