

以多層式社會網路為基之社交運算服務促進虛擬網路活動安  
全性之研究

計畫類別： 個別型計畫  整合型計畫

計畫編號： NSC 101-2221-E-343-001-MY2

執行期間： 101年 8 月 1 日至 103年 7 月 31日

執行機構及系所：南華大學電子商務管理學系

計畫主持人：陳宗義

共同主持人：陳裕民、陳垂呈

計畫參與人員：劉彥辰、沈筱娟、陳松奇、巫啟豪、蔡孟哲、許舒婷、黃瀨霆

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年  二年後可公開查詢

中 華 民 國 一 〇 三 年 九 月 十 日

## 一、中英文摘要及關鍵詞

### (一)計畫中文摘要

透過網路，一切實體世界的社交或商務活動已慢慢轉移延伸到無垠無界的虛擬網路世界。使得人際間的一切變得不再有距離，取而代之的是更多未知與風險。由於網際網路具跨時空、隱密、及匿名等特性，使人易脫於現實法律及道德的規範，造成網路亂象與網路活動風險不斷的提高，例如網路詐欺與金融犯罪等。為解決虛擬網路世界裡層出不窮的安全議題，若能夠利用一個人在網路上虛擬角色的活動紀錄進行社交運算(social computing)，建構一個兼具質與量的網路活動評估模式。例如，藉由網友相互的評價、虛擬社群發言、部落格的經營、及網路個人形象的營造來進行評估。藉此，使用者不論進行電子交易、網路交友或社群資訊分享等活動時，安全資訊能夠透過社交服務雲端技術隨處隨時可得，以提升網路虛擬活動的安全性，降低網路犯罪及風險。因此，本計畫將提出虛擬網路活動安全的整體解決方案，發展一以多層式的社會網路(multi-layer social network)為基之社交運算服務的機制，並輔以網路社交服務商務模式的設計，提供網路使用者安全社交服務，並以此為基發展可能的商業應用。為達上述目標，本研究主要的研究工作有：(1)研究領域與問題分析；(2)研究資料與需求收集；(3)以社交運算改善虛擬社群交友安全之方法設計；(4)推論網路使用者人格方法設計；(5)潛在顧客搜索方法設計；及(6)方法實作與測試。

**關鍵字：**社會網路、虛擬社群、信任評估、網路行為、潛在顧客。

### (二)計畫英文摘要

Through the Internet, almost all of real activities, such as friending and shopping, are transferred by degrees into and performed in virtual network worlds. Consequently, more and more unknown, unexpected events, and network risks occur from online social activities that have replaced distances between people has become a major issue. Due to the characteristics of the Internet such as unlimited time and space for online activities, privacy, and anonymity, the chaos and risks associated with the internet have resulted in network fraud and financial crimes. To solve these security issues, this study collects social activities from the Internet and constructs a social computing method that concerns quantifiable and qualitative data that evaluates the trust and reputation of network activities as virtual roles. When the evaluated information about trust and reputation to online virtual roles becomes readily available, the security of online social activities will be enhanced and will reduce network crimes and risks. This study will propose a total solution for virtual network activities including a multi-layer social network-based social computing service mechanism and a business model to offer secure services for network users. To achieve the objectives, this study will include: (1) studying issues related to this research topic; (2) collecting the data and requirements related to virtual network activities; (3) designing an approach based on social computing for improving the safety of making friend in virtual communities; (4) designing an approach for inferring network user's personality; (5) designing an approach for searching potential customer; and (6) implementing and testing the proposed approaches.

**Keywords:** social network, virtual community, trust evaluation, network behavior, potential customer.

## 二、計畫報告內容

### (一) 前言

在 Web 2.0 的時代，網際網路已成為共同合作的開放與分享的平台(洪淑民, 2010; 張乃文, 2009)，是一個新形態的人際互動空間(Shekarpour & Katebi, 2010)。Web 2.0 的網站可分成：以技術為主的工具型網站，具特定的功能或目的，如 Google 的搜尋引擎，提供使用者快又方便的網路資料搜尋服務；以社會網路為主的社群型網站，具有社群凝聚能力，例如 Facebook、MySpace、Friendster 及地圖日記等(洪淑民, 2010)，使用者藉由社群網站得以展現自己、連結社交網路，以及發展、維持與他人之關係。然而當彼此將虛擬世界的互動行為延伸至實體世界，由網路上熟悉的陌生人轉而成為實體世界具有實質關係的朋友時，當中可能潛藏未知的風險，例如網路一夜情、網路色情與詐騙等。

除了上述社會問題之外，虛擬社群同時也是企業關注的重要目標。由於虛擬社群獨具的開放性、互動性、及可連接網路上的同好，數以萬計的社群使用者，及其在社群網站上巨量的討論發言內容，更方便了企業進行目標行銷(targeted marketing)或可從中找尋潛在顧客(potential customer)。故以虛擬社交網路為基的電子商務更已成為未來商務模式及技術發展的趨勢(Lu et al., 2010)，買方與賣方透過虛擬社群進行產品資訊及口碑收集、討論並相互推薦，網路虛擬環境已完全融入我們真實的生活。未來雲端服務、行動商務與運算、及社交網路將匯聚成一個新的主流平台，雲端整合行動、行動整合社交、社交網路整合大量的資料與即時分析，已為市場迫切的需求。

因此，如何在虛擬的網路活動中利用社交運算(social computing)(Messinger et al., 2009; Kwai & Wagner, 2008)建構出有效測量成員網路活動品質及可被信任度之指標，找出虛擬分身網路互動中的關係，對電子商務、社群網站、虛擬社群成員等是相當重要的(劉麗萍, 2004; 連征忠, 2002; 陳仲偉, 2005)。若能善用網際網路之特性，將能夠解決許多虛擬網路活動所衍生的虛擬社會及交易問題，當然亦包含目前存在的一些實體社會問題，例如急難救助等。假如進一步藉由網路虛擬世界的虛擬化身的連結，進而構成實體世界的社會網路(social network, SN)，又在社交信任計算服務機制的防護下，虛擬社群使用者能夠安全的在進行線上娛樂、消費及社交活動。企業亦能藉此辨識潛在顧客及其顧客的人格特徵，致使企業能夠利用使用者於虛擬社群公開的討論內容，更有效精準的進行客製化行銷，提升企業獲利。因此，如何評估網路使用者於虛擬網路世界的社會活動品質及可信任度，及一套適當的虛擬網路分身評價的方法是本研究的主要目標。

### (二) 研究目的

本研究計畫藉由實體及虛擬社會網路模型的建構，使(1)在「實體方面」，能夠迅速連結實體使用者，機動的組織成虛擬團體於社會公益的推行，並應用有價值顧客的搜尋，整合應用於提升網路行銷的效益；(2)在「虛擬方面」，建構一針對網路虛擬活動的社交信任與信譽評估的服務模式，提升網路虛擬活動安全性以及降低網路犯罪及風險。為達上述目標，本計畫將設計「網路活動之使用者行為品質因素分析」、「虛擬社群互動安全理論架構」、及「虛擬社群互動安全評估方法」，並整合該產出應用於「潛在顧客搜尋模式」的發展，進而延伸設計「DISC 行為風格評估模式」的探討。再根據上述模式發展相關方法及技術，達到網路使用者網路行為品質及信任評估服務之計畫目標，進而發展其可能之商業服務及應用，使本研究更具學術及應用之價值。本計畫希望能提供安全的互動評估方式，讓虛擬社群使用者為自己做最安全的把關，以維護並創造個人的財產安全、人身安全、精神自由安全及個人資料安全。實體商業活動中，為企業或個人創造價值。

### (三) 文獻探討

由於本計畫研究工作項目繁多，因此需探究之相關理論及文獻者眾，這裡僅列舉整理「虛擬社群」、「資料探勘」、「潛在顧客」、「網路潛在風險」、「社交運算」、及「DISC 行為風格」等領域之文獻資料。

相關之理論及專有詞彙文獻，說明與整理如下：

● **虛擬社群**

隨著網際網路的發展，人際之間的交流已不再侷限於傳統的溝通方式。許多學者針對網際網路帶來的變革，依據不同的因素將網路使用者交流的空間加以定義，本研究參考 Lee 等學者(2002)所整理之虛擬社群定義，整理如表 3.1：

表 3.1：虛擬社群定義

定義	學者
虛擬社群是一個社會的集合體，透過人群帶著公開的議題、以及個人情感，來形成人際關係的網站。	(Rheingold, 1993)
虛擬社群是以電腦為媒介的，強調會員產生的內容、通訊等所整合而成的空間。	(Hagel & Armstrong, 1997)
虛擬社群是透明、開放的，允許團體、個人的參與，由以電腦為媒介的人際互動所促成的集合。	(Jones & Rafaeli, 2000)
人群透過電子媒體互相溝通，而非面對面。	(Romm & Clarke, 1995)
虛擬社群是一種社群的觀念，也就是說，它不能保證互動的機會，更確切的說，它必須自己產生互動。	(Craig & Zimring, 2000)
虛擬社群是一個拉長時間和地理限制的社群，由個人電腦和資訊的超級公路來補充現有建築與街道的不足。	(Hesse, 1995)
虛擬社群是一個長期、以電腦為媒介，大型團體之間的對話。	(Ericson, 1997)
虛擬社群是有關群眾的聚集。這些群眾因為提供了一個與其他人連結的環境，進行了一系列的營造信任氣氛、真正理解的互動，而因此被稱作虛擬社群。	(Carver, 1999)
虛擬社群是以科技為媒介，持續的存在、多元的互動風格、實際互動的功能與多使用者參與的環境。	(Ho et al., 2000)

根據表 3.1，本研究將虛擬社群之定義歸納為：「參與者使用電腦或其他電子媒體透過網際網路彼此互動、交換訊息之網際空間。」

● **資料探勘**

資料探勘是資料庫知識探索(Knowledge Discovery in Database)過程中的一個步驟。資料探勘係指自動搜尋隱藏在大量資料當中的特殊關聯與法則(Berry & Linoff, 1997)，而這些特殊關聯性可能隱含著過去未知且有價值的訊息，可提供決策者作為訂定決策的參考依據(Cabena et al., 1998; Han & Kamber, 2001)。隨著商業貿易的交易頻繁，資料探勘已廣泛被應用(Fayyad et al., 1996; Srikant & Agrawal, 1997; Sakakibara & Ohsawa, 2007; Hong, 2010; Chang, 2011)統計方法與資訊科技，可從大量的交易資料中，快速的萃取隱藏於資料中寶貝訊息。

資料探勘:的模型主要有四種：資料分類 (data classification)、資料關連 (data association)、資料分群 (data clustering) 以及循序樣式探勘 (sequential pattern mining)。常見的研究主題為關聯法則 (Association Rule)。探勘關聯法則過程最被廣泛使用係由Agrawal學者提出的Apriori演算法(Srikant & Agrawal, 1997)，其核心概念為找出最大項目集合。KeyGraph是由Ohsawa等人 (1998) 所創，屬於資料採礦 (data mining) 的一種演算法，將各種文章的觀點予以視覺化，在視覺化的過程中，將各個文章段落中經常同時出現的高頻單詞節點，和鄰近共同出現的單詞節點賦予有意義的連結，再解讀所產生的圖形做出結論。其特點有二，即視覺化圖形、和發現弱關係。

跨產業的資料探勘流程已被業界廣泛使用，此流程包含了可循環執行的：商業理解、資料理解、資料準備、建模、評估、部署等六個工作階段(Olson & Shi, 2007)。本研究將參考此資料探勘流程，建立於商務應用之Big Data探勘流程。另外，也將運用資料探勘技術分析Big Data中概念間之關聯，以獲取商品創新因子。

### ● 潛在顧客

目前尋找潛在顧客的方法大致可歸類為：(1) Marketing survey: 指以科學的方法收集市場資料，運用統計分析為企業管理者提供科學決策的訊息。American Marketing Association (AMA)定義市場調查為一種借由訊息，將消費者、顧客與大眾聯繫起來的活動，用於識別與界定市場營銷的機會與問題，產生改進和評價營銷活動，監控營銷績效、增進對營銷過程的理解(AMA, 2012)；(2) Data mining：是使用統計和數學方法在大量的資料集中，挖掘具有意義的關係、特徵和趨勢 (Han & Kamber, 2006; Ngai et al., 2009) 探討了可應用於CRM的各種data mining技術，包含decision tree, self-organizing maps, Markov chains model, k-means, data envelopment analysis, and logistic regression。探勘技術亦可從transaction data 及 web log 等尋找潛在顧客；(3) E-mail marketing：是利用電子郵件為其傳遞商業訊息到消費者的直銷形式；及(4) Opinion leader marketing：隨著 Web 2.0 崛起，消費者越來越重視部落客的文章內容，企業因此對意見領袖執行行銷術，達到推廣產品的效果(Sheng & Guergachi, 2008)。

### ● 網路潛在風險

在網路活動盛行，方便生活的同時，網路犯罪亦隨之崛起。虛擬社群使用者透過網路互動行為，相互分享、溝通、互動、及回饋。當彼此將虛擬世界的互動行為延伸至實體世界，由網路上熟悉的陌生人轉而成為實體世界具有實質關係的朋友時，當中可能潛藏未知的風險。

法務部(2011)分析犯罪狀況，發現除了傳統犯罪類型之外，因資訊科技發達，網路詐騙、網路購物詐欺、網路色情、網路恐嚇等犯罪案件有增加的趨勢，且透過虛擬網路的「隱匿性」及「複雜性」，致使犯罪事件的追查更具難度。新北市政府警察局永和分局(2012)歸納常見網路犯罪樣態，包含網路詐欺、網路恐嚇、網路誹謗、網路煽惑他人犯罪、網路色情、線上遊戲衍生犯罪、侵害著作權、網路上販賣毒品禁藥與網路駭客。

本研究整理網路及虛擬社群互動安全相關文獻(廖有祿與李相臣，2003)及新聞事件後，認為虛擬社群的網路互動行為中，承擔著四種潛在風險。以下就此四種風險、相關犯罪手法及案例提出說明：(1)財產風險：網路互動中，曾發生有心人士利用 MSN、即時通要求代買遊戲點數、代繳費用、匯款轉帳借錢、代收認證碼簡訊而利用受害者手機門號付費及其他金援請託之事項，對網路使用者進行錢財的詐騙。(2)人身風險：透過與網友見面或進行援交的過程，可能置人於危險當中而遭受強暴、暴力、殺害等涉及人身安全之情況。(3)精神自由風險：詐騙集團以感情或援交為誘餌，恐嚇取財；有時要求網路使用者寄送照片或打開視訊，加以錄製後成為威脅的工具；取得個資後，進行長期精神騷擾；以網路文字、圖像騷擾、或傳送個人訊息的網路霸凌行為。(4)個人資料風險：不肖人士假冒網路使用者，要求其他網友到部落格衝人氣，卻遇釣魚網站或偽裝檔名的超連結；或要求網友販賣公司個資，以圖自身利益。個資風險亦包括在購物消費過程中，因個資外洩而接獲詐騙電話之情形。

無數交友社群網站中，大量的真假資訊在虛擬網路世界中流竄，社群網站雖有要求使用者留下真實身份的規定及相關安全過濾措施，然而卻鮮少教導使用者如何在互動過程中或在見面之前，判斷來自各方的網友們。

### ● 社交運算

社交運算跨足計算科學與社會科學，以社會心理學(Social Psychology)、人機互動(Human-Computer Interaction)、社群網路分析(Social Network Analysis)、人類學(Anthropology)、組織

理論(Organization Theory)、社會學(Sociology)與計算理論(Computing Theory)為基礎。主要在探討如何透過電腦運算系統輔助人類的社會行為(Wang et al., 2007)。

社交運算是資訊系統的新研究領域，不僅改變了資訊傳遞與互動的方式，也讓使用者透過不同的裝置設備、以更動態的方式傳遞訊息(Parameswaran & Whinston, 2007)。社交運算讓使用者可以更容易分享創作、與人互動，亦能讓群體共同即時創作。以往，發佈新聞消息的權利，掌控在少數的媒體手中，一般人很難傳播分享屬於自己的訊息。而社交運算改變了世界的運行，運用社交運算平台，例如社群網站、部落格、電子郵件、即時通訊軟體、社交網路服務、維基服務、共享書籤等，任何網路使用者皆能自由發佈屬於自己的消息、思想或知識。社交運算以科技建立或重建個人社交習慣及社交內容，進而輔助社交。Parameswaran 與 Whinston (2007)認為社交運算平台具有幾項特色：有成功的品牌、能利用輕便的工具進行，具有動態資訊空間、客觀中立、品質取向、名譽式的管理。研究者觀察社交運算平台，亦發現多數人採用 Facebook 在平台上發表言論或新聞。

● DISC 行為風格

人格(Personality)，有時稱作性格、個性或行為風格，係指人類內在，具有一致性質的心理傾向與特徵。這種傾向與特徵在不同時間、地點與外在環境互動，影響一個個體的思考與行為(Wikipedia, 2013)。人格亦即個體獨特的外在行為，是人內在人格的體現(Bai et al., 2012)，Knapp & Daly (2002)提到由於個體人格的不同，造成溝通方式的差異。

自古至今，針對「人格分類」議題發展出多種分類理論。最早的分類法可追溯至古希臘時期之醫師—希波克拉底(Hippocrates)所提出的四體液學說，其認為四種體液之不平衡將導致人類不同的行為模式。四體液學說雖經後世證明此學說並無事實根據，然其偶然地提供一種認識人群的方法，往後發展出的各種分類理論，其源頭皆可追溯至此(MBAlib, 2013)。本研究項目採用之人格模型為美國心理學家 Moulton (1928)提出的 DISC 理論。DISC 理論將正常人分為四種行為風格類型，分別為：掌控型(Dominance)、誘導型(Inducement)、穩健型(Steadiness)、及服從型(Compliance)，此人格模型於廣泛地應用於現今的管理與商業用途上。

Boyd (1994)進一步改編了馬斯頓的模型，將四個分類以如圖 3.1 的四個象限的呈現方式。兩個垂直相交的軸分別代表了處事步調(Pace)與處事優先(Priority)。處事步調快者充滿自信、總是在力求改變；反之者往往較內斂及保守，處事以任務導向專注於要完成的事情，他們通常喜歡個人作業，較拘謹，並隱藏個人情緒。以人為導向者，處事則往往較注重與人的關係，對於既有的規定之態度，則較不拘束。四種行為風格者，具備有不同的特質，其一般描述經本研究整理後，呈現於圖 3.1。



圖 3.1：DISC 象限與一般描述圖

## ● 特徵選取(Feature Selection)

特徵選取(Feature Selection)常作為訓練分類器之前置作業，由於針對訓練資料所設定之原始特徵集之中，有些特徵可能是互有相依性或甚至是對於分類結果沒有影響力的，因此需要進行特徵選取，從原先的特徵集中挑選出更精確、更有鑑別力之有效特徵子集，以提升分類工作的正確率(Accuracy)並減少運算時間(Operation time)、樣本數量(Sample size)、與成本(Cost) (Yang & Honavar, 1998)。

若要取得「最佳」特徵子集，需窮舉所有可能的特徵子集並進行全面性測試(即暴力法)，特徵數量愈多則複雜度愈高，伴隨而來的高昂運算與時間成本是不可避免的。因此，各種取得「近似最佳」特徵子集而成本相對較低的方法相繼被提出。其方法可大致分為兩大類型：(1) Filter、與 (2) Wrapper。Filter 類型方法係透過衡量訓練資料自身特性的方式，例如計算亂度(熵，Entropy)、共同資訊量等統計指標，以評斷原始特徵是否夠資格作為特徵，此類型方法之計算量較 Wrapper 類型方法小，但得到的特徵子集品質較差；Wrapper 類型方法則是根據分類器結果，藉由不同的搜尋策略將產生高正確率的特徵納入特徵子集，雖相較前者計算量較大，但能得到較佳的特徵子集(國家教育研究院雙語詞彙、學術名詞暨辭書資訊網，2014)。

無論是採用哪種類型的方法進行特徵選取，其流程不外乎為以下四步驟(Liu & Motoda, 1998)：(1) 產生特徵子集，(2) 特徵子集的效能評估，(3) 判定是否達終止條件，及(4) 結果驗證。

本研究於特徵選取方法預計採用 Uysal & Gunal (2012) 所提出之核心概念，以特徵之條件機率為基對文件分類進行有效的特徵選取，此法屬於 Filter 類型的方法。因應用領域不同，本研究尚須依據應用領域對此法進行部分調整。

## (四)研究方法

為達本計畫之預定目標，本計畫即依下列之執行策略及工作方針，如下：(1) 創新研究與實務並重之研發；(2)分散式虛擬團隊之同步式研發；(3)跨計畫及領域的整合方式進行；及(4)以知識協作模式及知識管理技術執行本計畫。本計畫將運用同步工程及虛擬團隊之協同、協作與整合並進的方法與技術，並輔以專案管理及知識管理，實施本計畫之規劃、分工、進行、整合及研究成果的控管，使能發揮多功能團隊效率並達預期研究的效果與目標，以加速本計畫各階段各項任務的執行速度。本計畫之參與人員，每周舉行討論會，進行腦力激盪、心得交換、成果整合。

綜觀本計畫，執行過程中適逢商務環境的變遷與需求的改變，發現實際上有其困難之處。因此本計畫於計畫(I)第一年已做適當調整，計畫(II)第二、三年則大致依調整之結果執行之。本研究進度至日前為止，本計畫結果與調整後之計劃結果大致相符。本年度延續未完成之工作，重點在與其方法改善及文件投稿的準備工作。現階段，在已公開式資料為搜尋基礎的「虛擬世界實體人物搜尋機制」已有具體的方法被提出，以人物搜尋機制為主要應用的「潛在顧客搜索機制」亦已完成投稿文件的撰寫，另外在以社交運算改善虛擬社群互動安全的研究項目中，亦完成「虛擬社群互動安全評估方法的設計」。並已完成文件撰寫的工作。本計畫在依照預期的研究進度的執行過程中，發現推論網路使用者人格於行銷應用的可行性與重要性，因此在本年度的工作項目中，就原規劃進度外，特別將本計畫的研究項目延伸出一個子項目，進行利用使用者授權之公開於臉書的資訊，進行「推論網路使用者人格特質的方法」的研究。本年度成果報告，主要列舉本計畫下列三項研究項目與成果，概述如下：

- (1) 以社交運算改善虛擬社群交友安全之方法研究：本研究項目歸納虛擬社群網路安全互動決策因素提出「虛擬社群互動安全理論架構圖」，並自編「虛擬社群互動行為量表」測驗問卷，針對虛擬社群使用者不同的網路互動動機與目的，設計「休閒娛樂互動評估方法」、「消費互動評估方法」及「社交互動評估方法」，進行虛擬社群互動安全評估。以安全互動評估表產生各因素之得分，並利用評估公式產出安全評估值，提供虛擬社群使用者作為互動行為之參考。
- (2) 推論網路使用者人格方法研究：本研究項目提出「網路使用者人格推論方法」，本方法係基於此

項目另提出之「四階層社交媒體互動行為模型」與「預設特徵集」，及使用者的資料使用同意授權的基礎下，自動化地以網路使用者於社交媒體(本研究以 Facebook 為例)的互動行為紀錄，進行該使用者之人格推論工作。

- (3) 潛在顧客搜索方法研究：本研究項目鎖定網路上的使用者，並著重於中文語意處理，具體貢獻為提出「自動化領域詞彙淬煉方法」並提出「情緒字為基之向量空間模(Sentiment word based Vector Space Model, SVSM)」以及提出「虛擬社群中潛在顧客搜索模式」。

以上三項主要研究項目的研究細節與產出，及其所提出之方法與系統測試數據，分別詳細列於本研究報告後續章節。

#### (五)以社交運算改善虛擬社群交友安全之方法研究

本研究項目旨在找出虛擬社群使用者評估安全互動對象之決策因素，以該決策因素提出虛擬社群互動安全評估方法，藉由不同的評估構面，設計互動評估表，求出評估互動對象的安全評估值，以幫助廣大的虛擬社群使用者藉此評選安全的互動對象，使網路互動行為更有保障，避免不幸的社會事件發生。

#### 5.1 虛擬社群互動安全理論架構圖設計

本項研究工作首先設計虛擬社群互動安全因素理論架構(圖 5.1)，本項研究工作歸納虛擬社群網路安全互動決策因素，根據馬傳鎮(2008)的犯罪心理公式，分為個人層面、環境層面及情境層面。並依據 Goring (1913)犯罪心理學特質論，提出自我控制因素與自我概念因素；依據 Gottfredson & Hirschi (1990)一般犯罪理論，提出自我控制因素；依據 Hirschi (1969)社會控制理論提出依附因素、參與因素及信念因素；依據 Cohen & Felson (1979)日常活動理論，提出動機因素與機會因素；並由研究者的實際觀察中，提出信任因素、聲譽因素及時間因素，共計 10 項因素。以下分別說明其定義：

- (1) 自我概念(self-concept)：本研究將「自我概念」定義為虛擬社群使用者對自我整體的看法，包括自我外貌、自我評價等的描述。本研究以虛擬社群使用者觀察互動對象的網路社交名片及自述文字，視其對自我評價及自我概念是否偏低。若自我概念偏低者，涉入犯罪之可能性較高。
- (2) 自我控制(self-control)：本研究將「自我控制」定義為控制自我行為的能力。本研究中以虛擬社群使用者觀察互動對象的文章內容中，是否具有衝動(impulsive)、肉體物質(physical)、冒險(risk-taking)、投機(short-sighted)等特性為評估方法，若具有上述特性者，顯示其自我控制能力偏低，涉入犯罪之可能性較高。
- (3) 動機(motivation)：本研究將「動機」定義為虛擬社群使用者互動中所顯示出之行為意圖。研究方法中，以其意圖是否涉及財產、個資、身體裸露之請求為評估，若意圖涉及前述三者，則涉入犯罪之可能性較高。
- (4) 依附(attachment)：本研究定義「依附」為虛擬社群使用者對其在實體世界與虛擬世界社會網絡中的其他成員，包括家人、同儕及社群團體之情感的連結與附著。本研究以虛擬社群使用者觀察互動對象與親人、同儕、社群的連結數量或相簿數量為評估方法。其與真實世界之親人、同儕、社群或虛擬社群之連結數量或相簿數量愈低，則涉入犯罪之可能性較高。
- (5) 參與(involvement)：本研究定義「參與」為虛擬社群使用者加入正當社群活動、發展個人興趣、發表文章闡述其思想或分享個人經驗之程度。本研究以虛擬社群使用者觀察互動對象之發表文章數量、專業程度及所顯示之志趣為評估方法。其發表文章數量、專業程度愈低，且不具有顯著志趣者，則涉入犯罪之可能性較高。
- (6) 信念(belief)：本研究定義「信念」為虛擬社群使用者相信社會價值規範、倫理道德、信仰宗教、服從法律與法制之程度。本研究以虛擬社群使用者觀察互動對象之發表文章內容為評估方法。以文章內容中，顯示不符道德規範或缺乏信仰者涉入犯罪之可能性較高。
- (7) 信任(trust)：定義「信任」，其一為虛擬社群使用者對所屬虛擬社群或使用者進行互動之網站的信任；其二為虛擬社群使用者彼此互動中，累積安全互動經驗而產生的信任。以虛擬社群相識或互動場所是否為正當網站，及是否累積安全互動經驗、參與平台數量、及是否能被搜尋引擎印證其

真實度為評估方法。若互動場所為非正當網站，未累積充足之安全互動經驗，參與平台數量少，且不能以搜尋引擎印證其真實度者，則涉入危險之程度高。

- (8) 聲譽(reputation)：本研究將「聲譽」定義為虛擬社群使用者，在網路長期經營中，經由他人給予評價或口碑所建立的形象。並以正負評價、推薦數量、點閱人數、訂閱人數為評估方法。所獲得之正面評價、推薦數量、點閱人數、訂閱人數愈低，則聲譽差，涉入危險的程度愈高。
- (9) 機會(opportunity)：本研究認為抑制犯罪者不在場促成犯罪發生的機會，定義「機會」為虛擬社群使用者在互動中單獨面對之時機。以虛擬社群使用者與互動對象在實體世界之單獨會面為評估方法。若虛擬社群使用者之互動對象要求單獨見面，則涉入危險程度較高。
- (10) 時間(time)：本研究定義「時間」為虛擬社群使用者交往歷程之時間長度。本研究以虛擬社群使用者與互動對象之互動時間為評估方法。若互動時間較短，涉入危險程度較高。

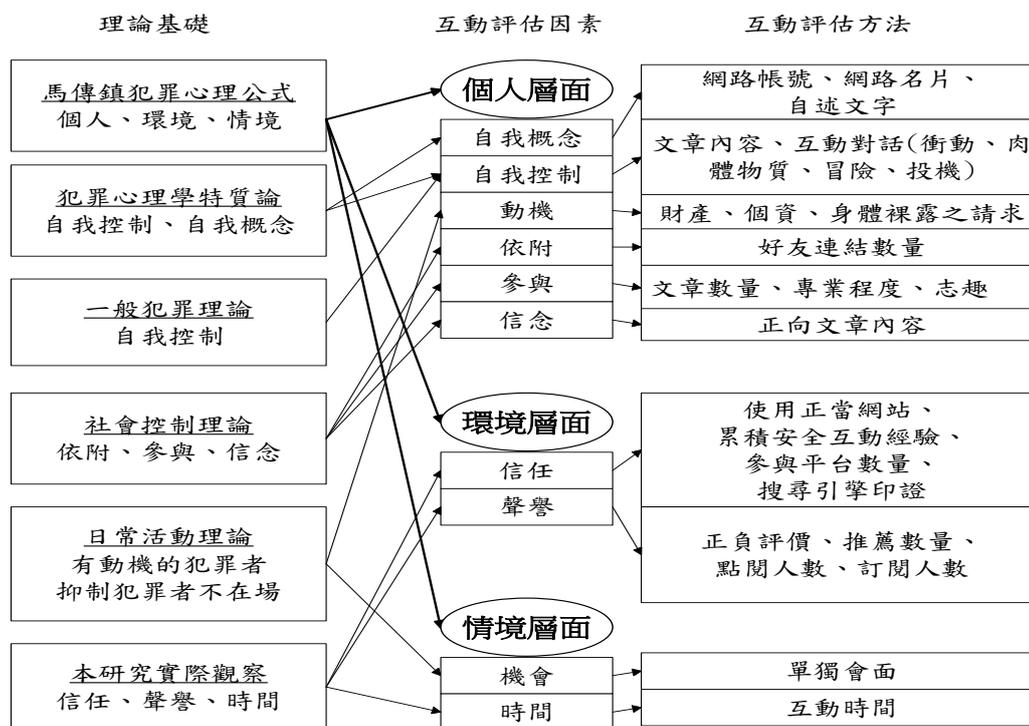


圖 5.1：虛擬社群互動安全因素理論架構圖

## 5.2 研究設計

本項研究子項目採用量化研究之問卷調查法進行研究測量，以台灣地區之虛擬社群使用者為研究對象，包括所有曾經加入虛擬社群並利用該社群進行互動行為之使用者。本子項目以圖 5.1 所提之各理論及安全因素為基礎，依據研究目的，自編而成「虛擬社群互動行為量表」為測量問卷。問卷內容由三個部分所組成：(1)網路互動經驗，探討虛擬社群使用者之實際互動經驗；(2)虛擬社群互動決策因素，測量虛擬社群使用者對於互動對象評估因素之同意程度；及(3)個人基本資料。本研究項目合併計算網路問卷及紙本問卷共 964 份，剔除填答者為非虛擬社群成員及填答不完整之無效問卷 206 份，共得有效資料 758 筆，可用率為 78.6%。本階段針對有效樣本進行敘述統計分析，其中包括：次數或人數分配、百分比分析、平均數以及變異數，藉以瞭解研究樣本的特徵與分布情形。

## 5.3 問卷結果與分析

本研究項目主要分析工作有：分析不同背景變項在網路互動經驗之差異，分析不同背景變項之虛擬社群成員在安全互動因素的認同差異，及分析虛擬社群互動安全因素之重要性分析。分析結果整理如下：

### ● 不同背景變項之虛擬社群在網路互動經驗的差異

本階段旨在了解虛擬社群使用者在網路互動行為的現況，與不同背景變項在此一層面的差異情形，

以完成研究目的之一：不同個人背景變項的虛擬社群使用者在網路互動行為的差異。本研究以社交活動、消費活動及休閒娛樂為主要研究範圍，以下分別就網路社交行為、網路消費行為及網路休閒娛樂行為進行統計分析。

- (1) **網路社交行為分析**：根據本研究項目所蒐集數據，在總樣本當中，36.1%的虛擬社群使用者在同一虛擬社群中同時擁有一個以上的帳號(274 人次)；55.7%的虛擬社群使用者曾在網路上和陌生網友短期或長期進行對談(422 人次)；8.0%的虛擬社群使用者曾經和陌生網友以電腦視訊方式見面(61 人次)；14.0%的虛擬社群使用者曾經和陌生網友在真實世界見面(106 人次)；32.7%的虛擬社群使用者曾經幫網友的部落格衝人氣(248 人次)。有 7 人曾經資助急需幫助的網友，佔 0.9%，有 751 人不曾資助陌生網友，佔 99.1%。在和陌生網友互動的時間長度統計中，46%的虛擬社群使用者不曾和陌生網友互動，其餘維持互動關係時間長度以一星期之內之 15.0%(114 人次)為最多，其次時間長度在一年內的有 10.8%(82 人次)，半年內的有 8.7%(66 人次)。並進一步進行交叉分析，如下：(a) 同時擁有一個以上帳號之網路行為：得知，在擁有多個帳號的網路行為中，以高中職學生最多，國中學生次之。可能原因為國高中學生之人格穩定度尚未趨於成熟，與其社交網路中的成員交往時，友誼關係較常變動；而當成員間的關係惡化或終止時，國高中學生除了將對方列為黑名單之外，往往會以新帳號重新開啟新的社交網絡關係，並將關係終止的成員排除在外。另外由國小階段升上國中，或由國中階段升上高中高職，因就學環境轉變，學生的同儕關係也同時更新。(b) 與陌生網友短期或長期進行對談或留言之網路行為：本研究的年齡樣本分層中，19~25 歲約為大學或大專學生，故推論，大學或大專學生有最多與陌生網友進行對談或留言之網路行為。(c) 與陌生網友進行視訊之網路行為：在與陌生網友進行視訊之網路行為中，男性遠多於女性。就年齡而言，集中於 31~40 歲，有 25 人次，佔 41.0%；其次為 19~25 歲，有 12 人次，佔 19.7%。以職業類別分析，學生族群為最多，有 26 人，佔 42.6%；其次為電子科技業，有 20 人，佔 32.8%。(d) 與陌生網友在實體世界中實際見面之網路行為：實際見面行為分析中只有 6 人，佔 5.7%。推論可能原因，在網路互動行為中，以文字訊息為溝通方式之想像空間最大，語音通話次之，視訊見面居三，而想像空間最小之溝通方式為實際見面。(e) 幫網友部落格衝人氣之網路行為：探討幫網友部落格衝人氣之網路行為，發現在曾經幫網友部落格衝人氣的 248 個樣本中，男性有 117 人，佔 47.2%；女性有 131 人，佔 51.2%。可知在曾經幫網友部落格衝人氣之網路行為中，女性多於男性。(f) 與陌生網友維持互動的時間長度：深入分析男女性在不曾互動的人數中，男性有 154 人，女性有 195 人，可見女性不曾與陌生網友互動的情形明顯偏多。
- (2) **網路消費行為分析**：本研究階段之受訪者 758 人中，73.4%(556 人)有網路消費經驗，26.6%(202 人)則不曾在網路上消費。在網路消費支付款項的方式部分，最多為選擇以便利商店代繳，有 460 人，佔 60.7%；其次為選擇線上刷卡，有 159 人，佔 21.0%；以 ATM 轉帳者有 141 人，佔 18.6%。常使用的網路消費類型中，以網路拍賣為最多，有 429 人，佔 56.6%；其次為虛擬網購，有 375 人，佔 49.5%；實體網購項目中有 197 人，佔 26.0%；網路市集有 183 人，佔 24.1%；在個體店家消費的有 138 人，佔 18.2%；以網路進行電視購物的有 117 人，佔 15.0%。在網路購物中，64 人曾經有詐騙經驗，佔 8.4%；694 人不曾有網路消費遭詐騙經驗，佔 91.6%。另有 7 人曾經透過未實際見面的網友進行投資，佔 0.9%，有 751 人不曾透過陌生網友進行投資，佔 99.1%。
- (3) **網路休閒娛樂行為分析**：在本研究階段樣本中，473 人有使用線上遊戲的經驗，佔 62.4%；有 285 人不曾使用線上遊戲，佔 37.6%。有 233 人曾經因為線上遊戲而產生交易行為，佔 30.7%；不曾因為線上遊戲而產生交易行為的有 525 人，佔 69.3%。在因線上遊戲而產生消費行為的交易對象中，有 211 人與遊戲公司交易，佔 27.8%；107 人對遊戲交易網站交易，佔 14.1%；有 186 人與一般玩家直接進行交易，佔 24.5%。478 人曾經有本人或親友帳號遭盜用的經驗，佔 63.1%；280 人不曾經歷帳號被盜用的事件，佔 36.9%。而當發生帳號遭盜用，或購物時個資外洩而被詐財的事件時，416 人會警告友人，佔 54.9%；307 人會向該網站投訴，佔 40.5%；31.8 人會向 165 反詐騙網站檢舉，佔 31.8%；237 人會到警局報案，佔 31.3%；105 人選擇不處理，佔 13.9%。經歷帳號被盜用而導致財物損失的事件後，有 399 人會繼續網路消費行為，佔 52.6%；359 人不敢再進行網路消費，佔 47.4%。

## ● 不同背景變項之虛擬社群在網路安全互動因素的差異

本階段主要在探討不同背景之虛擬社群使用者，對網路互動安全因素的認同程度上是否有差異。背景變項共分為性別、年齡、教育程度、職業、婚姻狀態、宗教信仰、手足排行等 7 項。本研究中，虛擬社群的安全互動因素共分為自我概念、自我控制、動機、依附、參與、信念、信任、聲譽、機會與時間等 10 項，採用獨立樣本  $t$  檢定與單因子變異數分析方法進行考驗，以下分別就不同背景變項說明之。

- (1) **不同性別之虛擬社群對網路安全互動因素的認同差異：**不同性別之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $t=5.198, p<.001$ )、自我控制( $t=6.152, p<.001$ )、動機( $t=6.033, p<.001$ )、信念( $t=5.797, p<.001$ )及機會( $t=6.915, p<.001$ )等五項因素中達.001 的顯著水準。意即不同性別之虛擬社群使用者，對自我概念、自我控制、動機、信念及機會等五項因素的認同程度呈現差異，且以平均數之數值分析，五項因素中，女性的認同程度皆高於男性，意即女性在評估網路互動安全因素時，比男性重視自我概念、自我控制、動機、信念及機會等因素。
- (2) **不同年齡之虛擬社群對網路安全互動因素的認同差異：**不同年齡之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $F=9.853, p<.001$ )、自我控制( $F=16.408, p<.001$ )、動機( $F=5.114, p<.001$ )、信念( $F=4.297, p<.001$ )、信任( $F=4.471, p<.001$ )及機會( $F=19.046, p<.001$ )等因素中達.001 的顯著水準。顯示不同年齡之虛擬社群使用者，對自我概念、自我控制、動機、信念、信任及機會等因素的認同程度呈現差異。進一步以 Scheffé 法進行事後多重比較，發現 31~40 歲、41~50 歲及 51 歲以上之組別較重視自我概念因素，亦即年齡層在 31 歲以上之受訪者較認同自我概念為安全互動之決策因素。自我控制因素中，19~25 歲的認同程度最低；動機因素中，41~50 歲的認同程度最高，而 13~15 歲的認同程度最低；信念因素中，51 歲以上的認同程度最高，而 13~15 歲的認同程度最低；信任因素中，19~25 歲的認同程度最高，而 13~15 歲與 16~18 歲的認同程度較低；機會因素中，19~25 歲的認同程度低於其他各組。
- (3) **不同教育程度之虛擬社群對網路安全互動因素的認同差異：**不同教育程度之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $F=19.244, p<.001$ )、自我控制( $F=12.536, p<.001$ )、動機( $F=9.904, p<.001$ )、信念( $F=11.281, p<.001$ )及機會( $F=11.622, p<.001$ )等因素中達 0.001 的顯著水準。顯示不同教育程度之虛擬社群使用者，對自我概念、自我控制、動機、信念、信任及機會等因素的認同程度呈現差異。進一步以 Scheffé 法進行事後多重比較，發現在自我概念因素中，研究所以以上及大學或大專的認同程度較高，而國中的認同程度最低；自我控制因素中，研究所以以上及小學的認同程度較高，而大學或大專的認同程度最低；動機因素中，研究所以以上的認同程度高於其他各組，而國中的認同程度最低；信念因素中，研究所以以上的認同程度最高，而的國中認同程度最低；機會因素中，研究所以以上的認同程度最高，而國中的認同程度最低。
- (4) **不同職業之虛擬社群對網路安全互動因素的認同差異：**不同教育程度之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $F=32.210, p<.001$ )、自我控制( $F=23.950, p<.001$ )、動機( $F=10.418, p<.001$ )、依附( $F=4.782, p<.001$ )、參與( $F=12.712, p<.001$ )、信念( $F=26.965, p<.001$ )、信任( $F=16.679, p<.001$ )、機會( $F=9.850, p<.001$ )及時間( $F=7.710, p<.001$ )等因素中達 0.001 的顯著水準。顯示不同職業之虛擬社群使用者，對自我概念、自我控制、動機、依附、參與、信念、信任、機會及時間等所有因素的認同程度皆呈現差異。進一步以 Scheffé 法進行事後多重比較，發現電子科技業之虛擬社群使用者，對自我概念、自我控制、動機、參與、信念、信任及時間等因素的認同程度皆比其它組別低，呈現極度明顯差異。推論其原因，本研究認為電子科技業之受訪者因在工作範圍中，常須對所研發之機密善盡保護職責，甚或簽署保密條款，不論對公司內部組織成員或非公司成員皆有防備之習性，故對各因素之認同產生苛刻效應，因而有極低之認同程度。
- (5) **不同婚姻狀態之虛擬社群對網路安全互動因素的認同差異：**不同婚姻狀態之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $t=7.719, p<.001$ )、自我控制( $t=6.695, p<.001$ )、動機( $t=4.334, p<.001$ )、信念( $t=4.461, p<.001$ )、聲譽( $t=3.285, p<.05$ )及機會( $t=6.161, p<.001$ )等六項因素中達.05 或.001 的顯著水準。意即不同婚姻狀態之虛擬社群使用者，對自我概念、自我控制、動機、信念、聲譽及機會等六項因素的認同程度呈現差異，且以平均數之數值分析，五項因

素中，已婚者的認同程度皆高於單身者。

- (6) 不同宗教信仰之虛擬社群對網路安全互動因素的認同差異：不同宗教信仰之虛擬社群使用者，對網路互動安全因素認同程度，在自我控制( $F=5.189, p<.001$ )、信念( $F=3.966, p<.001$ )、及機會( $F=2.975, p<.01$ )等因素中達.001 或.01 的顯著水準。然而進一步以 Scheffé 法進行事後多重比較，因 Scheffé 法較為嚴謹，比較後發現並無顯著差異。以有信仰與無信仰為組別，對網路互動安全因素認同程度進行獨立樣本  $t$  檢定。不同信仰之虛擬社群使用者，對網路互動安全因素認同程度，在自我概念( $t=1.971, p<.05$ )、自我控制( $t=4.487, p<.001$ )、信念( $t=2.437, p<.05$ )及機會( $t=3.853, p<.001$ )等四項因素中達.05 或.001 的顯著水準。意即有信仰與無信仰之虛擬社群使用者，對自我概念、自我控制、信念及機會等四項因素的認同程度呈現差異，且以平均數之數值分析，四項因素中，有信仰者的認同程度皆高於無信仰者。顯示有信仰者在評估網路互動對象時，較無信仰者重視自我概念、自我控制、信念及機會因素。歸納宗教信仰對安全互動因素之差異分析，發現不同宗教信仰者之間，對各因素認同程度並無顯著差異；而有信仰者與無信仰者對安全互動因素的認同程度，在自我概念、自我控制、信念及機會等四因素中有顯著差異。
- (7) 不同手足排行之虛擬社群對網路安全互動因素的認同差異：不同教育程度之虛擬社群使用者對網路互動安全因素認同程度，在自我概念、自我控制、動機、參與、信念及聲譽等因素中達顯著水準。然而進一步以 Scheffé 法進行事後多重比較，因 Scheffé 法較為嚴謹，比較後發現僅自我概念( $t=7.458, p<.001$ )與自我控制( $t=6.656, p<.001$ )呈現顯著差異。顯示不同手足排行之虛擬社群使用者，對自我概念與自我控制因素有不同的認同程度，且在兩因素中，排行老么的認同程度皆較低。

#### ● 虛擬社群安全互動因素之重要性分析

本階段分析虛擬社群使用者對網路安全互動因素的認同程度，以平均數表示虛擬社群使用者認為各因素之重要性，並加以排序。由於本研究採用 Likert 五點量表，其平均值為 3 分，因此受訪者填答分數超過此一平均值越多，表示虛擬社群使用者在網路安全互動因素的認同程度越高。

各因素之題項以平均數轉換計算後，得出 10 個因素之平均數及標準差，並依平均數之高低排序各因素之重要性。由得分最高至最低，分別為動機因素(4.49)、機會因素(4.20)、信念因素(4.12)、自我控制因素(3.57)、自我概念因素(3.55)、信任因素(3.47)、參與因素(3.36)、聲譽因素(3.34)、時間因素(3.26)、依附因素(2.92)。可知，虛擬社群使用者在網路互動安全評估中，對於動機、機會與信念三個因素的認同度較高；而對依附因素的認同度最低。本階段並以各因素的平均數作為虛擬社群安全互動評估方法中各因素之權重。

#### 5.4 虛擬社群互動安全評估方法設計

本研究子項目以虛擬社群使用網路活動之動機與目的為起點，以本研究架構中所提出之虛擬社群使用者安全互動決策因素模型為基礎，並參照決策因素與虛擬社群使用者實際網路行為之關連性，以社交運算為基礎，提出虛擬社群評選安全互動對象之方法。本項目以虛擬社群使用者在網路上的社交活動、消費活動與休閒娛樂活動為研究範圍，針對三種不同的活動類型，歸納其使用動機與目的(表 5.1)，分別適用各類的互動安全評估方法。

表 5.1：虛擬社群網路活動之動機與目的說明表

活動類型	動機與目的	參考文獻
社交活動	人際關係、尋找另一半、尋求志同道合的夥伴、喜歡到處參加虛擬社群活動、喜歡交朋友、交流心得、意見與生活點滴、搜尋資訊	Hagel & Armstrong (1997); Hamberger & Ben-Artzi (2000); 施建彬 (2007); 創市際(2010)
消費活動	商業消費、購物前搜尋參考意見	
休閒娛樂	打發時間、興趣、虛擬交易	

以本項目所提出動機、機會、信念、自我控制、自我概念、信任、參與、聲譽、時間及依附等 10 項因素之平均數為依據，做為虛擬社群互動安全評估公式中，各項因素之權重數值。因虛擬社群使用者對網路活動之不同動機與目的，本研究提出三種互動評估方法，分別為：虛擬社群網路休閒娛

樂互動安全評估法、虛擬社群網路消費互動安全評估法、及虛擬社群網路社交互動安全評估法。依據第三章的操作型定義及活動類型之動機與目的，為不同的活動類型決定適合的評估因素，並發展出適用的評估公式。其中，虛擬社群網路休閒娛樂互動安全評估法採用動機、信念、信任、聲譽及時間等 5 項評估因素；虛擬社群網路消費互動安全評估法採用動機、機會、信念、信任、參與、聲譽、時間及依附等 8 項評估因素；而虛擬社群網路社交互動安全評估法採用動機、機會、信念、自我控制、自我概念、信任、參與、聲譽、時間及依附等 10 項評估因素。

### 5.4.1 互動安全評估機制架構

本階段根據使用者可能需即時評估被評估者(evaluated user)的需求，設計一個互動安全評估機制功能架構 (圖 5.2)，本架構包含一個圖形化使用者介面(GUI)，及 5 個核心功能模組及 2 個資料庫，如下：

- (1) Questionnaire and User Behavior Data Collection Module：以線上問卷方式，對大量使用者及網路專家進行問卷調查，收集一般使用者的網路行為資料。
- (2) Data Analysis Module：對收集的網路使用者的人口統計變數進行統計分析，了解目前網路使用人口的基本狀況。
- (3) Crucial Factors and Weights Gaining Module：分析使用者對10因素在意程度的變化，並隨時調整權重質，使評估結果更客觀。
- (4) Evaluated User Questionnaire Module：提供三個評估量表，進行被評估者的線上評估，所填量表結果儲存於 Security User Case Base更參考。
- (5) Network User Behavior Quality Evaluation Module：為本系統的核心機制，以下一節所提的方法進行user security evaluation，產生一個三個向度的評估結果，供使用者參考。
- (6) Data Feedback and Updating Module：在評估為中高安全度以上者，使用者可以與被評估者進行進一步的網路活動，若為低度安全者，系統則會建議user立即停止進一步的互動。不論在實體或虛擬的互動之後，使用者可以進行資料的回饋，做為未來系統準確度及權重調整的依據。
- (7) Network User Behavior Database：用於儲存本系統對於預測的網路使用者及一些網路專家進行問卷的基本資料。
- (8) Security User Case Base：用於儲存被評估者的評估量表資料及於進行進一步互動過程中，曾發生危險的個案資料。

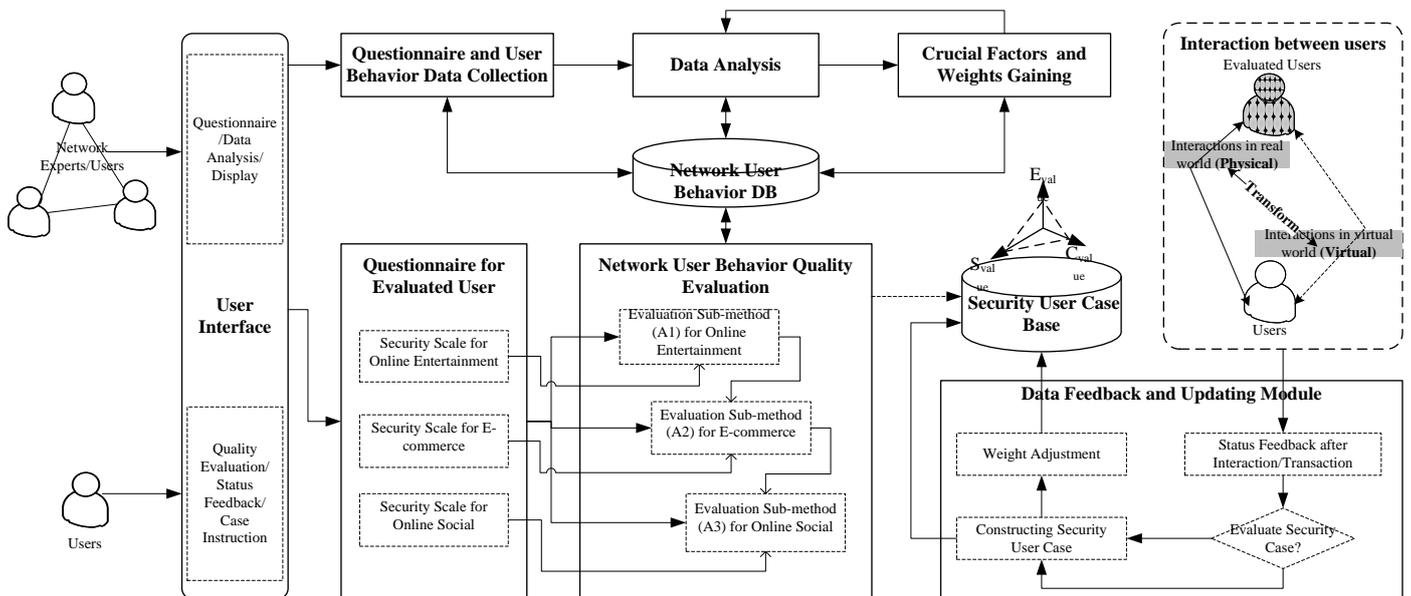


圖 5.2：互動安全評估機制功能架構

### 5.4.2 社交運算為基的互動安全評估法

本研究以虛擬社群使用者在網路上的娛樂、消費及社交活動為安全評估為範圍。針對這三種不同的活動類型，歸納其使用者的動機與目的，如下：(1)社交活動以人際關係的維持、尋求同好、交

朋友、及尋找另一半為目的；(2)消費活動以購物前搜尋參考意見，及商業消費為目的；及(3)休閒娛樂活動是以打發時間、興趣、及虛擬交易為目的。此三種活動類型可能依時間及互動雙方關係的深化，而相互移轉。例如，由休閒娛樂進入網路消費活動。

本研究子項目並且設計三個安全評估量表，分別為虛擬社群休閒娛樂互動安全評估表(表 5.2)、虛擬社群消費互動安全評估表(表 5.3)，與虛擬社群社交互動安全評估表(表 5.4)。供使用者於網路上與人互動時，需進行網路行為品質及安全評估時的參考。應用此三個量表，依評估者與評估對象之互動情境判斷，若符合評估方法之敘述，則選取「符合」欄中之分數為評估得分；若不符合評估方法之敘述，則選取「不符合」欄中之分數為評估得分，則可了解被評估者的網路行為品質為何。

**(1) 線上娛樂互動安全評估子方法**

線上娛樂互動安全評估子方法(ISE sub-method for online entertainment) 採用動機、信念、信任、聲譽、及時間等 5 項為評估因素。虛擬社群使用者進行休閒娛樂互動安全評估時，欲評估的對象可以是網友或遊戲平台，統稱為「被評估者」( $e_i$ )。本研究設計一個虛擬社群休閒娛樂互動安全評估表(表 5.2)，藉由本量表可求得被評估者的各評估項目得分。並將各得分導入函數(F3.1)，可求出安全評估值( $E_{value}$ )。

表 5.2：虛擬社群休閒娛樂互動安全評估表

因素	評估方法	符合	不符合	值域	評估得分
動機 ( $E_m$ )	網友要求我以手機代收認證簡訊( $E_{m1}$ )	0	1	{0, 1}	
	網友要求我幫他代買遊戲點數或儲值卡( $E_{m2}$ )	0	1	{0, 1}	
	網友要求我提供自己或他人的帳號密碼或個人資料( $E_{m3}$ )	0	1	{0, 1}	
信念 ( $E_b$ )	網路店家販賣非法遊戲光碟( $E_b$ )	0	4.29	{0, 4.29}	
信任 ( $E_u$ )	我在有認證機制的遊戲網站消費( $E_{u1}$ )	3.90	0	{0, 3.90}	
	我對該玩家或遊戲平台的安全消費經驗很多( $E_{u2}$ )	3.30	0	{0, 3.30}	
聲譽 ( $E_r$ )	玩家或遊戲平台的負評價很多( $E_{r1}$ )	0	3.74	{0, 3.74}	
	玩家或遊戲平台的推薦數量很多( $E_{r2}$ )	3.29	0	{0, 3.29}	
	玩家或遊戲平台的使用人數很多( $E_{r3}$ )	3.14	0	{0, 3.14}	
時間 ( $E_t$ )	玩家或遊戲平台已經經營很久了( $E_t$ )	3.46	0	{0, 3.46}	

$$E_{value}(e_i) = [w_b E_b(e_i) + w_{trust} E_u(e_i) + w_r E_r(e_i) + w_t E_t(e_i)] E_m(e_i) \quad , \quad (F5.1)$$

$E_{value}$ : 虛擬社群休閒娛樂互動安全評估值;  $0 \leq E_{value} \leq 88$  (四捨五入取整數);  $E_b$ : 信念因素得分;  $E_u$ : 信任因素得分;  $E_u = E_{u1} + E_{u2}$ ;  $E_r$ : 聲譽因素得分;  $E_r = E_{r1} + E_{r2} + E_{r3}$ ;  $E_t$ : 時間因素得分;  $E_m$ : 動機因素得分;  $E_m = E_{m1} \odot E_{m2} \odot E_{m3}$ , 即  $E_{m1}$ 、 $E_{m2}$  與  $E_{m3}$  當中, 只要有一者得分為 0, 則  $E_m$  的值即為 0, 否則為 1.

**(2) 線上消費互動安全評估子方法**

線上消費互動安全評估子方法(ISE sub-method for e-commerce)採用動機、機會、信念、信任、參與、聲譽、時間及依附等 8 項評估因素。User 進行消費互動安全評估時，欲評估的對象可以是賣家或購物網站( $e_i$ )。本研究設計虛擬社群消費互動安全評估表(表 5.3)，求得各評估項目得分。並將各評估得分導入函數(F5.2)中，可求出虛擬社群消費互動安全評估值( $C_{value}$ )。

表 5.3：虛擬社群消費互動安全評估表

因素	評估方法	符合	不符合	值域	評估得分
動機 ( $C_m$ )	賣家或購物網站來電要求我依照指示逐步操作 ATM 退款或轉帳( $C_{m1}$ )	0	1	{0, 1}	
	賣家或購物網站來電要求我核對個人資料( $C_{m2}$ )	0	1	{0, 1}	
機會 ( $C_o$ )	我將要單獨前往交貨地點與賣家碰面( $C_{o1}$ )	0	1	{0, 1}	
	賣家約我交貨的地點人煙稀少( $C_{o2}$ )	0	1	{0, 1}	
信念 ( $C_b$ )	網路店家販賣非法物品( $C_b$ )	0	4.29	{0, 4.29}	
信任 ( $C_u$ )	我在有認證機制的購物網站消費( $C_{u1}$ )	3.90	0	{0, 3.90}	
	我對該賣家或購物網站的安全消費經驗很多( $C_{u2}$ )	3.30	0	{0, 3.30}	
	賣家在許多不同的消費平台開設分店( $C_{u3}$ )	3.25	0	{0, 3.25}	
參與 ( $C_i$ )	某購物網站常與其他網站合辦促銷活動( $C_{i1}$ )	3.23	0	{0, 3.23}	
	賣家或購物網站能專業介紹或解說其商品( $C_{i2}$ )	3.32	0	{0, 3.32}	
聲譽 ( $C_r$ )	賣家或購物網站的負評價很多( $C_{r1}$ )	0	3.74	{0, 3.74}	
	賣家或購物網站的推薦數量很多( $C_{r2}$ )	3.29	0	{0, 3.29}	
	賣家或購物網站的點閱人數很多( $C_{r3}$ )	3.14	0	{0, 3.14}	
時間 ( $C_t$ )	賣家或購物網站已經經營很久了( $C_t$ )	3.46	0	{0, 3.46}	
依附 ( $C_a$ )	賣家或購物網站有開設實體店面( $C_a$ )	3.13	0	{0, 3.13}	

$$C_{value}(e_i) = [w_b C_b(e_i) + w_{trust} C_u(e_i) + w_i C_i(e_i) + w_r C_r(e_i) + w_t C_t(e_i) + w_a C_a(e_i)] C_m(e_i) C_o(e_i), \quad (F5.2)$$

$C_{value}$ ：虛擬社群消費互動安全評估值； $0 \leq C_{value} \leq 130$  (四捨五入取整數)； $C_b$ ：信念因素得分； $C_u$ ：信任因素得分； $C_u = C_{u1} + C_{u2} + C_{u3}$ ； $C_i$ ：參與因素得分； $C_i = C_{i1} + C_{i2}$ ； $C_r$ ：聲譽因素得分； $C_r = C_{r1} + C_{r2} + C_{r3}$ ； $C_t$ ：時間因素得分； $C_a$ ：依附因素得分； $C_m$ ：動機因素得分；在  $C_m$  的評估項目中， $C_{m1}$  與  $C_{m2}$  當中，只要有一者得分為 0，則  $C_m$  的值即為 0。意即虛擬社群使用者的情況只要符合其中一種敘述，則動機因素的得分即為 0。而只有在二者皆不符合的狀況中，動機因素  $C_m$  的得分為 1。 $C_o$ ：機會因素得分；在  $C_o$  的評估項目中， $C_{o1}$  與  $C_{o2}$  當中，只要有一者得分為 0，則  $C_o$  的值即為 0。意即虛擬社群使用者的情況只要符合其中一種敘述，則機會因素的得分即為 0。而只有在二者皆不符合的狀況中，機會因素  $C_o$  的得分為 1。

### (3) 線上社交互動安全評估子方法

線上社交互動安全評估子方法(ISE sub-method for online social)採計全部 10 項評估因素。虛擬社群使用者進行社交互動安全評估時，欲評估的是交友對象( $e_i$ )。利用虛擬社群社交互動安全評估表(表

5.4)，求得各評估項目得分，導入函數(F5.3)即可求出虛擬社群社交互動安全評估值 ( $S_{value}$ )。

表 5.4：虛擬社群社交互動安全評估表

因素	評估方法	符合	不符合	值域	評估得分
動機 ( $S_m$ )	網友要求我提供或出借金錢( $S_{m1}$ )	0	1	{0, 1}	
	網友要求我以手機代收認證簡訊( $S_{m2}$ )	0	1	{0, 1}	
	與網友視訊時，對方要求我裸露身體( $S_{m3}$ )	0	1	{0, 1}	
	網友要求我寄送我的裸露照片( $S_{m4}$ )	0	1	{0, 1}	
	網友要求我提供或買賣自己或他人的個人資料( $S_{m5}$ )	0	1	{0, 1}	
機會 ( $S_o$ )	剛認識的網友約我單獨見面( $S_{o1}$ )	0	1	{0, 1}	
	網友約我見面的地點人煙稀少( $S_{o2}$ )	0	1	{0, 1}	
信念 ( $S_b$ )	網友總是張貼違反倫理道德的事( $S_{b1}$ )	0	4.09	{0, 4.09}	
	網友常教人如何鑽法律的漏洞( $S_{b2}$ )	0	3.99	{0, 3.99}	
自我控制 ( $S_l$ )	網友與我的對話呈現衝動性格( $S_{l1}$ )	0	3.78	{0, 3.78}	
	從網友的文章內容看出他很重視物質或身體享受( $S_{l2}$ )	0	3.56	{0, 3.56}	
	網友喜歡冒險刺激的活動( $S_{l3}$ )	0	3.17	{0, 3.17}	
	網友的文章經常提到他的投機行為( $S_{l4}$ )	0	3.78	{0, 3.78}	
自我概念 ( $S_c$ )	網友的社交名片中使用負面詞語自我介紹( $S_{c1}$ )	0	3.75	{0, 3.75}	
	網友使用貶低自我的暱稱與帳號( $S_{c2}$ )	0	3.65	{0, 3.65}	
	網友的文章中總認為自己與別人的互動關係很差( $S_{c3}$ )	0	3.36	{0, 3.36}	
	網友文章內容中，常覺得自己能力差、沒有價值( $S_{c4}$ )	0	3.43	{0, 3.43}	
信任 ( $S_u$ )	我在色情網站或非正當網站認識該網友( $S_{u1}$ )	0	3.90	{0, 3.90}	
	我與網友的安全互動經驗多( $S_{u2}$ )	3.30	0	{0, 3.30}	
	網友以相同帳號參與許多不同的網路平台( $S_{u3}$ )	3.25	0	{0, 3.25}	
	從網路搜尋引擎查詢到關於網友本身的事實資料，與網友提供的資訊相符( $S_{u4}$ )	3.43	0	{0, 3.43}	

參與 ( $S_i$ )	網友的文章常分享他參加正當活動的經驗( $S_{i1}$ )	3.53	0	{0, 3.53}
	網友對特定領域的專業度很高( $S_{i2}$ )	3.32	0	{0, 3.32}
	網友有明顯的個人興趣抱負( $S_{i3}$ )	3.38	0	{0, 3.38}
聲譽 ( $S_r$ )	網友的負評價很多( $S_{r1}$ )	0	3.74	{0, 3.74}
	網友的推薦數量很多( $S_{r2}$ )	3.29	0	{0, 3.29}
	網友的點閱人數很多( $S_{r3}$ )	3.14	0	{0, 3.14}
	網友的部落格文章有很多人訂閱( $S_{r4}$ )	3.18	0	{0, 3.18}
時間 ( $S_t$ )	我跟網友互動的時間已經超過1年了( $S_t$ )	3.19	0	{0, 3.19}
依附 ( $S_a$ )	網友的網路相簿裡有他和家人同事的合照( $S_a$ )	3.13	0	{0, 3.13}

$$S_{value}(e_i) = [w_b S_b(e_i) + w_{s-cont} S_l(e_i) + w_{s-c} S_c(e_i) + w_{trust} S_u(e_i) + w_i S_i(e_i) + w_r S_r(e_i) + w_t S_t(e_i) + w_a S_a(e_i)] S_m(e_i) S_o(e_i), \quad (F5.3)$$

$S_{value}$ : 虛擬社群社交互動安全評估值;  $0 \leq S_{value} \leq 281$  (四捨五入取整數);  $S_b$ : 信念因素得分;  $S_b = S_{b1} + S_{b2}$ ;  $S_l$ : 自我控制因素得分;  $S_l = S_{l1} + S_{l2} + S_{l3} + S_{l4}$ ;  $S_c$ : 自我概念因素得分;  $S_c = S_{c1} + S_{c2} + S_{c3} + S_{c4}$ ;  $S_u$ : 信任因素得分;  $S_u = S_{u1} + S_{u2} + S_{u3} + S_{u4}$ ;  $S_i$ : 參與因素得分;  $S_i = S_{i1} + S_{i2} + S_{i3}$ ;  $S_r$ : 聲譽因素得分;  $S_r = S_{r1} + S_{r2} + S_{r3} + S_{r4}$ ;  $S_t$ : 時間因素得分;  $S_a$ : 依附因素得分;  $S_m$ : 動機因素得分; 在動機因素  $S_m$  的評估項目中,  $S_{m1}$ 、 $S_{m2}$ 、 $S_{m3}$ 、 $S_{m4}$  與  $S_{m5}$  當中, 只要有一者得分為 0, 則  $S_m$  的值即為 0。意即虛擬社群使用者的情況只要符合其中一種敘述, 則動機因素的得分即為 0。而只有在五者皆不符合的狀況中, 動機因素  $S_m$  的得分為 1。 $S_o$ : 機會因素得分; 在  $S_o$  的評估項目中,  $S_{o1}$  與  $S_{o2}$  當中, 只要有一者得分為 0, 則  $S_o$  的值即為 0。意即虛擬社群使用者的情況只要符合其中一種敘述, 則機會因素的得分即為 0。而只有在二者皆不符合的狀況中, 機會因素  $S_o$  的得分為 1。

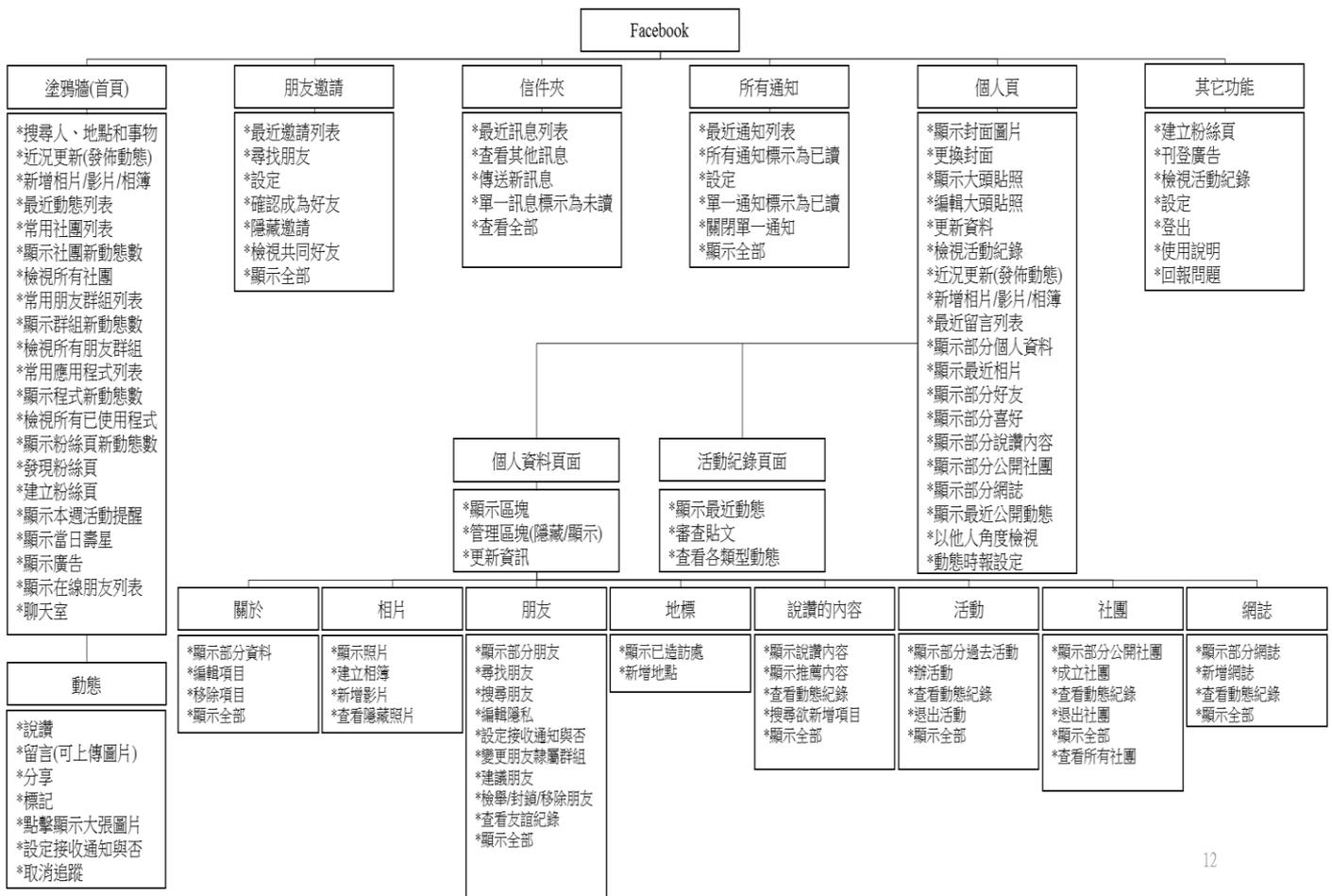
本項目提出之三個評估方法, 進行可行性分析採立意抽樣, 尋訪虛擬社群使用者共 50 人進行試驗評估。獲得具有休閒娛樂動機之有效樣本 19 筆; 具有網路消費動機之有效樣本 22 筆; 具有網路社交動機之有效樣本 15 筆。並以其四分位數分析以上資料, 區分為 4 個安全層次, 包含: 低度安全(0~Quartile1)、中低度安全(Quartile1~Quartile2)、中高度安全(Quartile2~Quartile3)、及高度安全(Quartile3~), 以此做為評估參考的常模。結果顯示: (1) 休閒娛樂安全評估結果中, 有 6 人屬低度安全, 建議立即停止互動行為; 3 人屬中低度安全, 建議審慎評估後, 再決定是否繼續互動行為; 分別各有 5 人屬中高度安全及高度安全, 建議可安心繼續互動行為。(2) 網路消費安全評估中, 有 5 人屬低度安全, 建議立即停止消費行為; 5 人屬中低度安全, 建議審慎評估後, 再決定是否繼續消費行為; 分別有 7 人及 5 人屬中高度安全及高度安全, 建議可繼續消費行為。(3) 網路社交安全評估中, 有 4 人屬低度安全, 建議立即停止交友行為; 1 人屬中低度安全, 建議審慎評估後, 再決定是否繼續交友行為; 分別有 6 人及 4 人屬中高度安全及高度安全, 建議可繼續交友行為。

## (六) 推論網路使用者人格

本研究項目提出一利用網路使用者公開於社交媒體(以 Facebook 為例)之互動行為的歷程紀錄, 作為推論其人格類型的依據。此研究項目所採用之人格模型(即 DISC 理論)中, 人格類型稱為「行為風格類型」, 然為免混淆, 往後皆統稱「人格類型」。以下依序介紹「社交媒體互動行為模型」、「特徵設計」與「行為編碼」等本研究項目提出之方法所使用的元素, 並於 6.4 節詳述方法之細部程序。

### 6.1 社交媒體互動行為模型

本研究以當前世界最大之社交媒體 Facebook (Wikipedia, 2014) 為社交媒體之實例, Facebook 之目前所提供之功能架構如圖 6.1 所示。



12

圖 6.1：Facebook 功能架構

根據觀察與整理，使用者於此社交媒體之互動行為結構，絕大多數為「使用者對其他使用者的某些物件進行動作」，例如：「Alice likes Bill's photo」或是「Cathy shared Denny's link」，諸如此類。

承上，圖 6.2 所示為本研究基於對使用者在 Facebook 上的互動行為之觀察所建立的互動行為模型。此模型共可分為四階層，第一階層為「行為者層(Actor-tier)」，代表觀察對象；第二階層為「行為層(Behavior-tier)」，代表行為者於 Facebook 上所進行的互動行為操作；第三階層為「目標對象層(Target user-tier)」，代表行為者進行互動行為操作所針對的對象；及第四階層為「目標物件層(Target object-tier)」，代表行為者進行互動行為操作所針對之物件，由於物件無法單獨存在，必有一目標對象作為其所有者，因此此階層係位於目標對象層之下。其中「行為層」、「目標對象層」、「目標物件層」在 Facebook 此一社交媒體中，各含有其可能的實例元素，亦一併呈現於圖 6.2 中，各實例元素在 Facebook 上代表的意義以及其代碼(6.2 節與 6.3 節將用上)於表 6.1 中說明。

基於此規律的行為模式，本研究項目參照前人之相關研究(Golbeck et al., 2011; Adalı & Golbeck, 2012; Adalı et al., 2012; Ortigosa et al., 2014)，設計了套用於 Facebook 互動行為之預設特徵集，此特徵集將應用於人格推論方法，特徵詳述於 6.2 節，人格推論方法則詳述於 6.4 節。

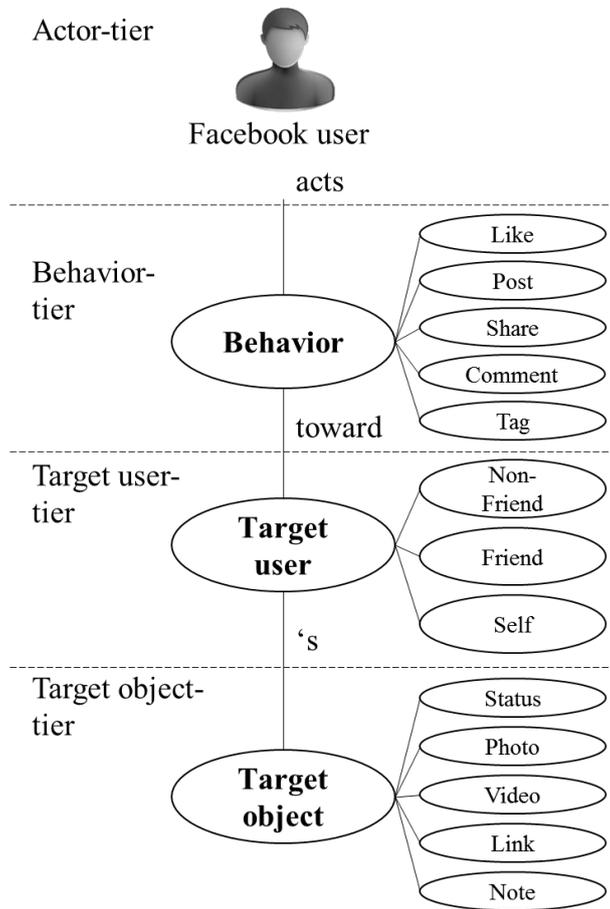


圖 6.2：四階層社交媒體個人單次互動行為模型-以 Facebook 為例

表 6.2：實例之代碼與中文說明對應表

類別	實例	代碼	中文說明
Behavior	Like	L	按讚
	Post	P	發表
	Share	S	分享
	Comment	C	留言
	Tag	T	標註
Target user	Non-friend	N	非好友名單中的對象
	Friend	F	好友名單中的對象
	Self	S	自己本身
Target object	Status	S	近況動態
	Photo	P	照片
	Video	V	影片
	Link	L	連結
	Note	N	網誌

## 6.2 特徵設計

本階段設計之互動行為預設特徵高達上百個，其具有共同之邏輯結構，其結構模型於圖 6.3 以樹狀圖來表示。根節點的 User A 代表觀察對象(即圖 6.2 中的行為者)，根節點以外之其他節點為特徵

分類，以下再進一步詳述各分類代表意義。

對於任一 Facebook 使用者，本研究針對該使用者於 Facebook 上進行的五種常見互動行為分別為 Like、Post、Share、Comment 與 Tag 進行特徵設計；於每一種行為底下再針對行為作用的目標對象設計了四種類型的分類，分別為 Global、Target User (TU)、Target Object (TO)、Target User-Object (T-UO)。

Global 類型的特徵只針對「當前行為」而不論目標對象或物件進行統計，例如：Like/Global/Count/N 分類中，有一特徵代碼為 N\_Like，其意義便為 Like 此一動作的總數；TU 類型的特徵只針對「當前行為作用的目標對象」進行統計，例如：Like/TU/Count/N 分類中，有一特徵代碼為 N\_LNf，其意義為對非好友 Like 的總數；TO 類型的特徵只針對「當前行為作用的目標物件」進行統計，例如：Comment/TO/Count/N 分類中，有一特徵代碼為 N\_CPt，其意義為對照片的 Comment 總數；T-UO 類型的特徵則針對當前行為作用的「目標對象複合物件」進行統計，例如：Share/T-UO/Count/N 分類中有一特徵代碼為 N\_SNfLk，其意義為對非好友之連結的 Share 總數。每一種行為類型都有兩種統計標的類型，一種為統計數量類型(Count)，此類型底下有總數量特徵(N)、平均數量特徵(AvgN)與數量亂度特徵(EtpN)；另一種為統計時間類型(Time)，此類型底下有最大間隔時間特徵(MaxT)、最小間隔時間特徵(MinT)、平均間隔時間特徵(AvgT)與間隔時間亂度特徵(EtpT)。部分特徵實例可見表 6.2。

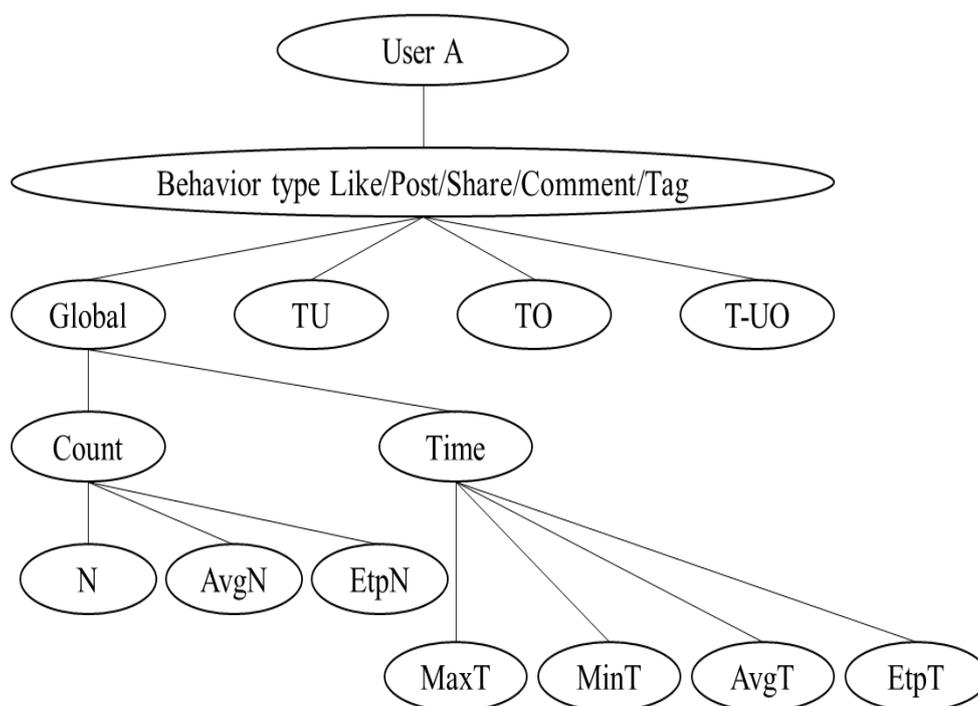


圖 6.3：互動行為特徵結構模型

表 6.2：預設特徵集之部分

Behavior type	Target type	Count type	Abbreviation type	Description	Abbreviation		
Like	Global	count	N	The number of <b>Likes</b> A made.	N_Like		
				The number of <b>individual user</b> A made <b>Likes</b> .	N_UL		
			AvgN	Average of <b>Likes</b> A made for <b>individual users</b> . ( $N\_Like/N\_UL$ )	AvgN_L		
			EtpN	Entropy of number of <b>Likes</b> A made to <b>individual users</b> .	EtpN_LU		
		time	MaxT	Maximum time interval between adjacent <b>Likes</b> A made.	MaxT_L		
			MinT	Minimum time interval between adjacent <b>Likes</b> A made.	MinT_L		
			AvgT	Average time interval of adjacent <b>Likes</b> A made.	AvgT_L		
			EtpT	Entropy of time interval between adjacent <b>Likes</b> A made.	EtpT_L		
			Target User, TU	count	N	The number of <b>Likes</b> A made to <b>non-friends</b> .	N_LNf
						The number of <b>Likes</b> A made to <b>friends</b> .	N_LF
	The number of <b>Likes</b> A made to <b>self</b> .	N_LSf					
	The number of <b>non-friends</b> A made <b>Likes</b> .	N_NfL					
	The number of <b>friends</b> A made <b>Likes</b> . ( $N\_F$ : the number of all A's friends)	N_FL					
	AvgN	Average of <b>Likes</b> A made for <b>non-friend</b> . ( $N\_LNf/N\_NfL$ )			AvgN_LNfL		
		Average of <b>Likes</b> A made for <b>all A's friend</b> . ( $N\_LF/N\_F$ )			AvgN_LF		
		Average of <b>Likes</b> A made for <b>A's Liked friend</b> . ( $N\_LF/N\_FL$ )			AvgN_LFL		
	EtpN	Entropy of number of Likes A made to non-friends.			EtpN_LNfL		
		Entropy of number of Likes A made to friends.			EtpN_LFL		
	time	MaxT		Maximum time interval between adjacent <b>Likes</b> A made to <b>non-friends</b> .	MaxT_LNfL		
				Maximum time interval between adjacent <b>Likes</b> A made to <b>friends</b> .	MaxT_LFL		
				Maximum time interval between adjacent <b>Likes</b> A made to <b>self</b> .	MaxT_LSf		
		MinT		Minimum time interval between adjacent <b>Likes</b> A made to <b>non-friends</b> .	MinT_LNfL		
				Minimum time interval between adjacent <b>Likes</b> A made to <b>friends</b> .	MinT_LFL		
				Minimum time interval between adjacent <b>Likes</b> A made to <b>self</b> .	MinT_LSf		
		AvgT		Average time interval between adjacent <b>Likes</b> A made to <b>non-friends</b> .	AvgT_LNfL		
				Average time interval between adjacent <b>Likes</b> A made to <b>friends</b> .	AvgT_LF		
				Average time interval between adjacent <b>Likes</b> A made to <b>self</b> .	AvgT_LSf		
		EtpT		Entropy of time interval between adjacent <b>Likes</b> A made to <b>non-friends</b> .	EtpT_LNfL		
			Entropy of time interval between adjacent <b>Likes</b> A made to <b>friends</b> .	EtpT_LF			
			Entropy of time interval between adjacent <b>Likes</b> A made to <b>self</b> .	EtpT_LSf			
Target Object, TO		count	N	The number of <b>Likes</b> A made to <b>statuses</b> .	N_LSt		
				The number of <b>Likes</b> A made to <b>photos</b> .	N_LPt		
				The number of <b>Likes</b> A made to <b>videos</b> .	N_LVd		
	The number of <b>Likes</b> A made to <b>links</b> .			N_LLk			
	The number of <b>Likes</b> A made to <b>notes</b> .			N_LNt			
	The number of <b>statuses</b> A made <b>Likes</b> .			N_StL			
	The number of <b>photos</b> A made <b>Likes</b> .			N_PtL			
	The number of <b>videos</b> A made <b>Likes</b> .			N_VdL			
	The number of <b>links</b> A made <b>Likes</b> .			N_LkL			
	The number of <b>notes</b> A made <b>Likes</b> .			N_NtL			

### 6.3 互動行為編碼

本節介紹互動行為編碼，互動行為編碼係基於 6.1 節互動行為模型之實例，將使用者於社交媒體上之「互動行為紀錄」轉為「代碼字串」之過程。6.2 節之表 6.1 已呈現互動行為模型中各實例與其代碼對照，以下表 6.3 則呈現相鄰互動行為間之間隔時間區間(Time interval)與其代碼對照。

表 6.3：間隔時間區間代碼表

類別	實例	代碼
Time interval	0~1 hr.	1
	1~4 hr.	2
	4~8 hr.	3
	8~12 hr.	4
	12~16 hr.	5
	16~20 hr.	6
	20~24 hr.	7
	24 hr. up	8

參照 6.1 節之互動行為模型，與表 6.1 與表 6.3 二代碼對照表，可將 Facebook 中任何互動行為紀錄轉換為一字串( $R$ )， $R$  為一系列的  $r_1, r_2, \dots, r_n$  等單次互動行為與相鄰行為之間隔時間( $t_i$ )所串接成的長字串。其符號定義如下：

$$R = r_1 t_1 r_2 t_2 r_3 t_3 \dots t_{n-1} r_n ;$$

$r_i$ ：表單次互動行為， $r_i = buo$ ；；

$b$ ：表行為類型， $b \in \Sigma_B, \Sigma_B = (L, P, S, C, T)$ ；

$u$ ：表目標對象類型， $u \in \Sigma_U, \Sigma_U = (N, F, S)$ ；

$o$ ：表目標物件類型， $o \in \Sigma_O, \Sigma_O = (S, P, V, L, N)$ ；及

$t_i$ ：表相鄰行為之間隔時間， $t \in \Sigma_T, \Sigma_T = (1, 2, 3, 4, 5, 6, 7, 8)$ 。

以下針對  $r_i$  進行編碼之動作，列出三個範例與說明：

- (1) Amy likes Bill's photo 之編碼為 LFP。假設某一使用者 Amy 對於某一個在其好友名單內的好友 (Friend) Bill 所發布的照片(Photo)按讚(Like)，此動作可被編碼成為 LFP 並成為 Amy 之互動行為紀錄編碼的一部份。
- (2) Amy shared Cathy's link 之編碼為 SNL。Amy 分享了某一不在其好友名單內的使用者(Non-friend) Cathy 所分享的連結(Link)，此動作可被編碼成為 SNL，並成為 Amy 之互動行為紀錄編碼的一部份。
- (3) Amy commented on her own status 之編碼為 CSS。Amy 在她自己(Self)的近況動態(Status)底下留言(Comment)，此動作可被編碼成為 CSS，並成為 Amy 之互動行為紀錄編碼的一部份。

若範例中三次的行為是在某時間區間內依序發生，並分別間隔 30 分鐘與 6 小時(間隔時間區間代碼請參照表 6.3)，則 Amy 在此時間區間內的互動行為編碼字串即為： $R = LFP1SNL3CSS$ 。

透過此編碼方式可將一時間區間內的互動行為紀錄轉為一長字串，如此便可針對大量相同人格類型者之互動行為紀錄編碼，從中挖掘高頻率出現之相似字串，並以挖掘得出的相似字串作為該人格類型的代表成分，進而比對未知人格類型者的互動行為紀錄中含有各人格類型之代表成份多寡來判定該使用者之人格類型。

### 6.4 推論網路使用者人格類型方法設計

本研究項目提出二種人格推論方法，分別為「特徵分類法」與「成分比對法」，分述如下。

#### ● 特徵分類法

為一般分類常用方法，分為訓練與應用兩階段，圖 6.4 為此特徵分類法之流程圖，實際流程分為四步驟，描述如下：

- (1) 蒐集足量訓練資料：在此步驟須先建立一人格測驗網頁，測驗內容依使用之人格模型而定，例如本研究採用 DISC 行為風格作為人格模型，則可使用 <<THE PLATINUM RULE: Discover the Four Basic Business Personalities and How They Can Lead You to Success>> (Alessandra & O'Connor, 1998)書中所提之 DISC 測驗量表。將此測驗網頁作為 Facebook APP 型態，Facebook 使用者可於 Facebook 框架下讀取此網頁內容，並於進入網頁作答前，需經由 Facebook 提示視窗要求使用者，同意授權此 APP 日後可透過 Facebook API 合法取得其個人資料。此即意味，當使用者完成測驗，本系統即可紀錄其測驗分類結果，並同時透過測驗網頁 APP 取得該使用者資料(本研究將存取由使用者同意授權使用的個人「動態時報(Timeline)」的歷程資料)。
- (2) 篩選關鍵特徵：並非所有的預設互動行為特徵，最後都能夠對人格分類有所貢獻，因此需從中選取適當的特徵集用於訓練分類器，此即特徵篩選。此步驟需先對各筆訓練資料計算其預設特徵值，並計算各特徵與分類結果的相關係數(Pearson coefficient)，最後僅保留特徵值與分類結果高度相關者，當作關鍵特徵。
- (3) 訓練分類器：利用步驟(2)所篩選得出的關鍵特徵集來訓練分類器。至於採取何種分類器較為恰當，本研究項目預計於後續實作與實驗中進行測試，並做成建議。
- (4) 進行人格分類與校正分類器：前三步驟皆屬於「訓練階段」，第四步驟則屬於「應用階段」。本項目預期於應用階段建立全新的人格分類 Facebook APP，單純採集經授權之未知人格類型的實驗參與者之互動行為紀錄，並運用步驟(3)所訓練之分類器進行人格分類，呈現分類結果予受試使用者同時亦提供使用者評估結果準確性之功能，藉由實際分類結果與使用者反饋結果，定期校正分類器。

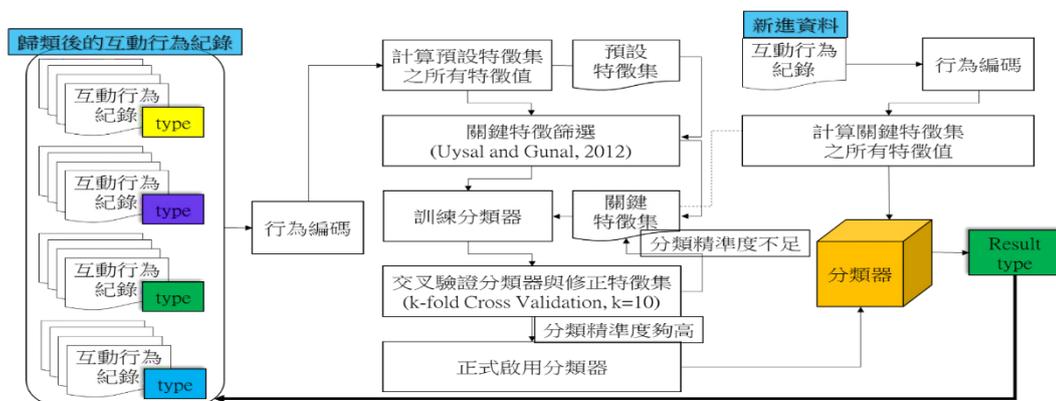


圖 6.4：特徵分類法流程圖

## ● 成分比對法

本方法包含訓練與應用兩階段如圖 6.5 所示，成分比對法與前法不同之處在於本方法不需訓練分類器，也毋須事先設計大量預設特徵，僅需了解目標社交媒體之互動元素(請參照 6.2 節之模型)。本法之前置作業為「行為編碼」，請參閱 6.3 節進行對照。此方法步驟如下：

- (1) 蒐集足量訓練資料：與前法之步驟(1)相同，可共用由 APP 蒐集來的同一資料集(即已標註之互動行為紀錄)。
- (2) 互動行為編碼：將各筆訓練資料依照 6.3 節之規則進行編碼，每筆訓練資料將成為一長字串。
- (3) 代表成分挖掘：所謂「代表成分」係指於標註為同一人格類型的所有互動行為記錄中經常出現的相似行為序列，其為在各互動行為編碼字串裡的子字串。此步驟利用本研究所提出之 Top-k SPM

演算法，可挖掘出各人格類型之代表成分。

- (4) 設定成分權重：步驟(3)可找出各人格類型之代表成分集，但代表成分之長度與出現頻率必定不一。此步驟之任務，便是依據代表成分長度與出現頻率為每一代表成分設定權重，作為往後新進互動行為紀錄資料計算與各人格類型相似度之依據。權重設定與計算相似度分數之實作方法尚未制定，原則上，一筆新進資料與某一人格類型之相似度分數，必然與該新進資料包含有此人格類型代表成分之「數量」、包含成分之「長度」以及包含成分出現於同一人格類型者之「頻率」成正相關。
- (5) 進行人格分類與校正成分權重：本步驟可使用與前法步驟(4)相同之人格分類 APP，並採用步驟(3)挖掘出的代表成分，以及步驟(4)為各成分設定的權重為依據，對新進資料進行人格類型相似度計算與分類。呈現分類結果予受試使用者同時亦提供使用者評估結果準確性之功能，藉由實際分類結果與使用者反饋結果，定期校正成分權重。

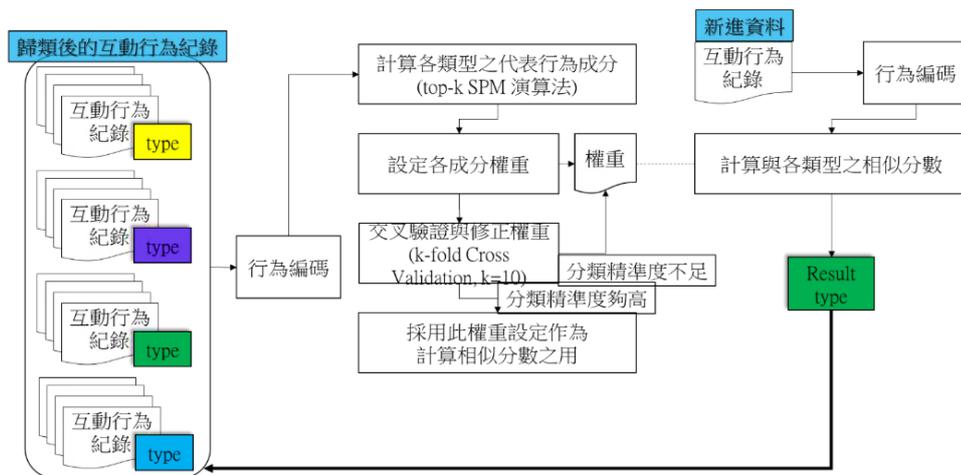


圖 6.5：成分比對法流程图

綜合以上二法，可將本網路使用者人格推論系統呈現於圖 6.6。特徵分類法之方法流程為一般分類議題經常使用，成分比對法則為本研究創新提出。兩方法之基本概念截然不同，因此本研究預期將於實驗階段時採用以上二種方法並比較分類精準度，且致力於提升本研究提出的法二之精準度。

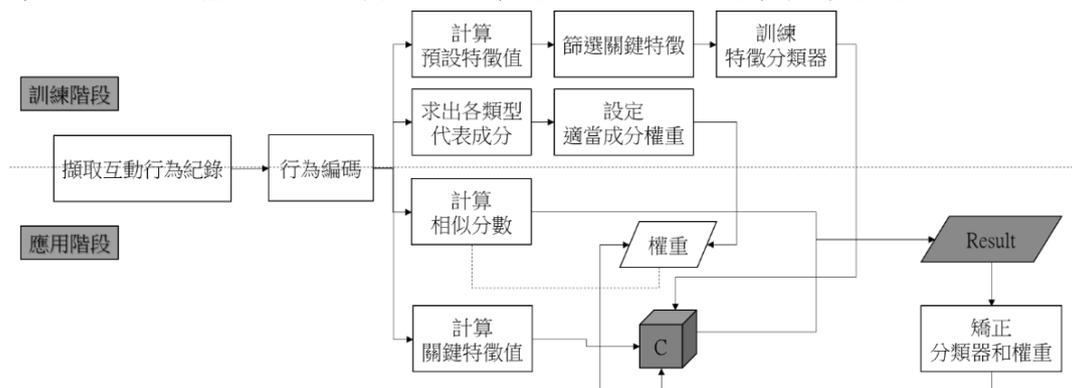


圖 6.6：推論網路使用者人格類型系統運作流程

### (七) 潛在顧客搜索

由前項研究項目的調查結果得知，極高比例的受訪者有網路消費經驗及使用虛擬社群的習慣，因此本研究子項目認為若由企業角度切入，利用本計畫研究之初所提之自動化實體人物之網路公開資訊搜尋機制之「虛擬世界實體人物公開資訊搜尋模式」及「人物輪廓模型」等方法，在虛擬社群之公開的討論區上搜尋特定企業之潛在顧客。因此本項目提出潛在顧客搜索方法，協助本計畫所得之結果能夠適用於網路行銷活動，精準的為業者找到適當的顧客。本階段首先提出一個潛在顧客搜索方法，其

流程主要活動及方法之細部分別說明於下面子節。

### 7.1 領域辭庫建構

為精確搜尋潛在顧客資訊，使用者必須先給予欲進行潛在顧客搜索之企業之基本輪廓資訊，以便建立與企業及其產品相關之詞庫。因此本研究設計企業資訊輪廓模型(Enterprise Profile Model, 圖 5.1)，包含：(1)企業，欲分析潛在顧客的公司資訊；(2)其他企業，包含產業中合作與競爭的公司資料；(3)商品類別，該公司之所屬產業及商品類型；(4)商品，該公司生產或販售之所有商品；(5)原料及組件，產品之成分及主要零組件；(6)產品特性，描述產品主要特徵及規格；(7)事件，企業近年來所發生的重大事紀；(8)服務，所能提供的服務項目等。

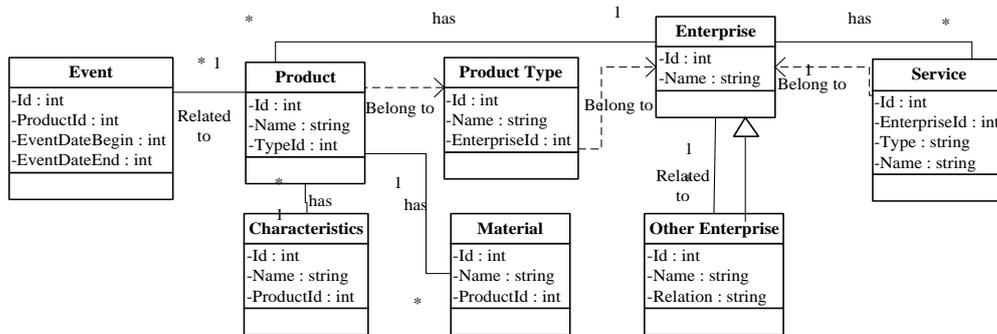


圖 7.1：企業資訊輪廓模型

本階段設計之辭庫建構流程(Lexicon Construction Process)，由於詞性標註服務(Part Of Speech Tagging Services, POS Tagging Service)並無法正確識別專有名詞，例如「奶油酥餅」一詞經由詞性標註(POS Tagging)後會分為「奶油(Na)」及「酥餅(Na)」兩個詞彙。上述專有名詞係由兩個以上的名詞所組成，本研究將此種關連定義為連續名詞(Continuity Noun)之關係。為避免在關鍵字擴展階段搜尋相關元素時，領域詞彙誤斷所造成的困擾，本研究設計一以搜尋引擎為基之字詞合併機率計算公式，利用共現(Co-occurrence)關係計算字詞為連續名詞之可能性，提出可識別誤斷之詞彙，並重新合併連續名詞之演算法。流程如下：

- (1) Query Merging：藉由企業輸入之資訊，以產品名稱作為關鍵字，將各關鍵字與該產業常見之一般情緒詞(Common Sentiment word)合併，合併順序預設為關鍵字在前，情緒詞在後，兩詞中間以空白分隔。關鍵字擺放順序對多數搜尋引擎而言不影響其搜尋結果。然而搜尋引擎會將文字拆解進行相似搜尋，為避免搜尋引擎拆解文字後模糊搜尋結果，可利用搜尋引擎提供的自訂選項進行輸入，以 Google 為例，以雙引號包住的文字在搜尋結果的網頁中必須完全符合才會回傳。
- (2) Searching：透過步驟(1)所合併之查詢，發送至搜尋引擎中查詢其共現程度。一般搜尋引擎在搜尋時會重新導向(Redirect)至另一串網址，係以「查詢字串(Query String)」之方式將參數(Parameter)列於網址中(含關鍵字)，並未採用相對嚴格的門票法，可自訂參數訪問特定網址，因此可藉由程式指定欲搜尋之關鍵字進行自動化查詢。
- (3) Domain Term Refinement：此步驟之目的為獲取各關鍵字之誤斷合併率，並取產品名稱中最低之誤斷合併率供後續特殊詞彙淬煉(Specific Term Refinement)判斷之門檻值(Threshold)。誤斷合併率係指專有名詞在詞性標註服務中所造成之字串錯誤切割之情形，透過本研究所提出的方法，以此合併率為門檻值，將誤斷之字串重新合併為專有名詞，演算法如圖 7.2 所示。

Line	Algorithm	Introduction
1 :	<b>Input</b> : product name array $P$ , combination rate array $C$ .	
2 :	<b>Output</b> : refinement threshold $T$ , base word of a specific term $Base$ .	
3 :	<b>For</b> $i$ from 1 to Count( $P$ ) <b>do</b>	分別計算領域詞彙( $P$ )之誤斷合併率.
4 :	$t \leftarrow \text{SendToCKIP}(P(i));$	將第 $i$ 個領域詞彙送至詞性標註服務斷詞.
5 :	<b>If</b> (Count( $t$ )>1) <b>Then</b>	若專有名詞被斷開，進行 Line 6-11.
6 :	$cnt \leftarrow 1$	
7 :	<b>For</b> $j$ from Count( $t$ ) to 1 <b>do</b>	將斷開之詞彙陣列反轉後填入陣列 $k$ .
8 :	$k(cnt) \leftarrow t(j)$	
9 :	$cnt++$	
10 :	<b>End For</b>	
11 :	push last_word( $t$ ) into $Base$	將斷開詞彙中最後一截單詞存入陣列 $Base$ .
12 :	<b>Else</b>	若專有名詞並未被斷開，標註誤斷合併率為 1，並跳至迴圈的下一詞彙.
13 :	$C(i) \leftarrow 1$	
14 :	<b>Continue For</b>	
15 :	<b>End If</b>	
16 :	$X \leftarrow \log(\text{GetSearchResult}(k(1), \dots, n))$	將 $k(1), \dots, n$ 之查詢結果數量存入變數 $X$ .
17 :	$Y \leftarrow \log(\text{GetSearchResult}(k(1), \dots, n-1))$	將 $k(1), \dots, n-1$ 之查詢結果數量存入變數 $Y$ .
18 :	$Z \leftarrow \log(\text{GetSearchResult}(k(n)))$	將 $k(n)$ 之查詢結果數量存入變數 $Z$ .
19 :	$C(i) \leftarrow X / \text{sqrt}(Y*Z)$	以 Eq. 1 進行計算，得到合併機率 $C(i)$ .
20 :	<b>End For</b>	
21 :	$T \leftarrow \text{Median}(C)$	取 $C$ 之中位數，作為門檻值回傳.
22 :	<b>Return</b> $T$	
		$C = \frac{\log F[k(1, \dots, n)]}{\sqrt{\log F[k(1, \dots, n-1)] \times \log F[k(n)]}}$ $n \in N^+ \quad (F7.1)$

圖 7.2 : Domain Term Refinement Algorithm

## 7.2 關鍵字擴展

經由前一節之程序，得到已初始化之特殊詞彙淬煉之門檻值，接著透過本節之關鍵字擴展機制(Keyword extension mechanism)，以產品名稱作為基礎詞彙，將與產業相關之關鍵字，透過網際網路進行搜索：

- (1) Content Retrieval：關鍵字擴展之目的在於，將與產業有關聯之專有名詞萃取出來，存入詞庫。而輸入之關鍵字，將做為種子(Seed)用於搜尋與其相關之文章，再從文章中尋找專有名詞。然而產品名稱本身較具識別度，搜索出來的結果通常與企業相關，導致難以擴展出新的字詞。因此本研究藉由搜索產品原料或產品零組件，將有機會取得更多與產業相關之專有名詞，如：提及奶油，可能還會提及紅豆、車輪餅等專有名詞，其餘之細節同辭庫建購流程之 Searching 步驟。
- (2) Content Pre-Processing：此步驟保留不同格式(Format)中之純文字內容，並將其以詞性為單位進行詞性標註。
- (3) Specific Term Refinement：如圖 7.3 的演算法所示，目的在於尋找新詞彙，若利用詞性進行搜尋，會存在以下問題：(1)具有同一詞性之詞彙過多，(2)專有名詞誤斷，及(3)無法判斷同詞性之名詞是否為該領域中的詞彙。因此，本項目利用 F7.1 獲得特殊詞彙淬煉之門檻值。

Line	Algorithm	Introduction
1 :	<b>Input</b> : sentence array with POS tagging $S$ , product name array with POS tagging $P$ , threshold $T$ , base word of a	
2 :	specific term $Base$	
3 :	<b>Output</b> : specific term array $Specific$	
4 :	$Specific \leftarrow$ <b>new</b> Array	初始化 $Specific$ 為一個空的陣列.
5 :	<b>For</b> $i$ from 1 to count( $S$ ) <b>do</b>	分別將每個已斷開之語句逐字尋找特殊詞彙.
6 :	$cnt \leftarrow 1$	
7 :	<b>For</b> $j$ from Count( $S(i)$ ) to 1 <b>do</b>	將斷開之詞彙陣列反轉後填入陣列 $k$ .
8 :	$k(cnt) \leftarrow S(i)(j)$	
9 :	$cnt++$	
10 :	<b>End For</b>	
11 :	$cnt \leftarrow 1$	
12 :	$baseCnt \leftarrow 1$	
13 :	$begin \leftarrow$ <b>False</b>	
14 :	<b>For</b> $j$ from 1 to count( $k$ ) <b>do</b>	將 $k$ 逐項進行特殊詞彙識別.
15 :	<b>If</b> ( $k(j)$ exists in $Base$ ) <b>Then</b>	
16 :	$baseCnt \leftarrow j$	若 $k$ 中第 $j$ 項字詞符存在於基礎詞陣列中, 則設定起始數
17 :	$begin \leftarrow$ <b>True</b>	$baseCnt$ 為 $j$ , 並設定開始狀態 $begin$ 為真(true).
18 :	<b>End If</b>	
19 :	<b>If</b> ( $begin$ and $j > baseCnt$ ) <b>Then</b>	若開始狀態為真, 且 $j$ 大於起始數則:
20 :	$X \leftarrow$ GetSearchResult( $k(baseCnt, \dots, j)$ )	將 $k(1, \dots, n)$ 之查詢結果數量存入 $X$ .
21 :	$Y \leftarrow$ GetSearchResult( $k(baseCnt, \dots, j-1)$ )	將 $k(1, \dots, n-1)$ 之查詢結果數量存入 $Y$ .
22 :	$Z \leftarrow$ GetSearchResult( $k(j)$ )	將 $k(n)$ 之查詢結果數量存入 $Z$ .
23 :	$R(cnt) \leftarrow \log(X) / \text{sqrt}(\log(Y) * \log(Z))$	以 Equation 1 代入計算之, 獲得 $R$ .
24 :	<b>If</b> ( $R(cnt) \geq T$ ) <b>Then</b>	若合併機率大於等於門檻值, 則累加 $cnt$ .
25 :	$cnt++$	
26 :	<b>Else If</b> ( $cnt > 1$ ) <b>Then</b>	若 $R$ 低於門檻值, 但 $cnt$ 大於 1, 則將 $k(baseCnt, \dots, j-1)$
27 :	push $k(baseCnt, \dots, j-1)$ into $Specific$	存入 $Specific$ 陣列中.
28 :	<b>Else</b>	若合併機率低於門檻值, 且 $cnt$ 小於等於 1, 將 $begin$ 、 $cnt$ 、
29 :	$begin \leftarrow$ <b>False</b>	$R$ 等變數初始化, 進行下個迴圈.
30 :	$cnt \leftarrow 1$	
31 :	clear( $R$ )	
32 :	<b>End If</b>	
33 :	<b>End If</b>	
34 :	<b>End For</b>	
35 :	<b>End For</b>	
36 :	<b>Return</b> $Specific$	

圖 7.3 : Special Term Refinement Algorithm

(4) Term Filtering: 步驟(3)之結果包含已標註之詞性, 本步驟依據該服務之定義(CKIP service, 2012),

針對普通名詞(Na)、專有名詞(Nb)以及領域特殊詞彙(Specific)進行篩選，若在已篩選資料庫中存在該名詞，則略過該名詞之處理；若未找到該名詞則視為新詞彙，進行相似度計算(Term Similarity Calculation)。

- (5) Term Similarity Calculation：相似度計算係基於 Vector Space Model (VSM)(Salton et al., 1975)進行改良。本研究項目設計一個情緒字為基之向量空間模型 Sentiment word based Vector Space Model (SVSM)，藉由重新定義 VSM 中的維度，用來計算任兩名詞間之關聯程度。本研究項目將 VSM 之相關公式重新定義如下：

$$s_i = (P_{1,i}, P_{2,i}, \dots, P_{t,i}) \quad t, i \in N^+ \quad (F7.2)$$

$$j = (P_{1,j}, P_{2,j}, \dots, P_{t,j}) \quad t, j \in N^+ \quad (F7.3)$$

$$\cos \theta = \frac{s_i \cdot j}{\|s_i\| \|j\|} \quad (F7.4)$$

$$0^\circ \leq \theta \leq 180^\circ$$

其中， $s_i$ ：種子名詞  $i$  之共現頻率向量； $P_{t,i}$ ：文件集中，種子名詞  $i$  與第  $t$  個情緒詞之共現次數； $j$ ：目標名詞之共現頻率向量； $P_{t,j}$ ：件集中，目標名詞  $j$  與第  $t$  個情緒詞之共現次數； $s_i \cdot j$ ：向量  $s_i$  與向量  $j$  之向量內積； $\|s_i\|$ ：量  $s_i$  的映射長度； $\|j\|$ ：向量  $j$  的映射長度； $\cos \theta$ ：種子名詞  $i$  與目標名詞  $j$  的向量餘弦值。

F7.2~7.4 將 VSM 之維度重新定義為情緒詞(Sentiment word)與名詞在文件集合中之 co-occurrence。此一概念假設每個名詞皆有許多用以形容該名詞之情緒詞彙，而在文件集合中，兩個毫無關連的名詞在情緒詞彙相同的情況下，其 co-occurrence 會有顯著差異，即夾角偏差較大；而有高度關聯性的兩個名詞其夾角偏差較小。

由於適合描述每個名詞的情緒詞皆有所不同，查詢任兩名詞之關聯，必須以相同的情緒詞來比較，因此以下陳述將兩名詞分為「種子(Seed)名詞」與「目標(Target)名詞」來區分，種子名詞為詞庫中原有之專有名詞；而目標名詞則為系統偵測到不存在於資料庫的專有名詞。

由於文件集中，名詞與情緒詞之間的文字距離(Text Distance)若較遠，可能代表該情緒詞並非用於形容此一名詞，因此本研究將 F5.4 展開後，加入文字距離中位數作為兩者關聯度的計算，如 F7.5 所示。

$$Similarity(i, j) = \cos \theta = \frac{\sum_{t=1}^n \left( \frac{P_{t,i}}{MedianDist_{t,i}} \times \frac{P_{t,j}}{MedianDist_{t,j}} \right)}{\sqrt{\sum_{t=1}^n \left( \frac{P_{t,i}}{MedianDist_{t,i}} \right)^2} \times \sqrt{\sum_{t=1}^n \left( \frac{P_{t,j}}{MedianDist_{t,j}} \right)^2}} \quad (F7.5)$$

$$0^\circ \leq \theta \leq 180^\circ, n \geq 1 \wedge n \in N^+$$

其中， $P_{t,i}$  為文件集中，種子(Seed)名詞  $i$  與第  $t$  個情緒詞之共現次數； $P_{t,j}$  為文件集中，目標(Target)名詞  $j$  與第  $t$  個情緒詞之共現次數； $MedianDist_{t,i}$  為種子名詞  $i$  與第  $t$  個情緒詞之文字距離

中位數;  $MedianDist_{i,j}$  為目標名詞  $j$  與第  $t$  個情緒詞之文字距離中位數;  $\cos \theta$  為種子名詞  $i$  與目標名詞  $j$  的向量餘弦值;  $Similarity(i, j)$  為種子名詞  $i$  與目標名詞  $j$  的相似度, 越接近 1 越相似。

上述方法之文件集並不侷限於正規文件(Regular Document), 名詞與情緒詞亦不侷限於正式格式(Regular Format). 在本研究中, 文件集採用搜尋引擎之資料庫、名詞與情緒詞之標註採用(CKIP service, 2012)之標註結果, 名詞與情緒詞之共現次數為在搜尋引擎中查詢兩者共同出現的社群網頁總數。

### 7.3 潛在顧客篩選

本階段透過潛在語意分析(Latent Semantic Analysis, LSA)方法分析網路文章。潛在語意分析屬於向量空間模型的延伸, 由 Deerwester et al. (1990)所提出, 主要應用於解決一字多義(Polysemy)、同義字(Synonym)的問題。LSA 是一種以數學為基礎的知識模型, 利用奇異值分解(Singular Value Decomposition, SVD)來找出字詞對應文件的語意結構(Semantic Structure), 並可透過維度約化(Dimension Reduction)過濾雜訊, 並保留原始矩陣之特徵(Deerwester et al., 1990)。LSA 將文件與詞彙之間的關聯表示為  $t \times d$  的矩陣  $X$ , 其中每一個元素表示詞彙  $t$  在文件  $d$  的出現次數。  $X$  經由 SVD 分解後, 得到三個矩陣  $T$ 、 $S$  與  $D$ , 其中矩陣  $S$  為一對角矩陣, 代表語意空間,  $T$  代表的是詞彙在語意空間中的表示法,  $D$  則代表文件在語意空間的表示法。透過 dimension reduction 重建之矩陣表達為  $X \sim X' = T'S'D'^T$ 。對一未知文件  $D_q$  (Pseudo-document), 可透過 F7.6 計算該文件於語意空間(Semantic Space)  $S'$  中之向量, 再以 Cosine 相似度計算其與  $D'^T$  中各文件之相似度。

$$D_q = X_q'^T T' S'^{-1} \quad (F7.6)$$

其中,  $X_q'^T$  為各字詞在  $D_q$  中出現次數之轉置矩陣(Transpose Matrix);  $T'$  為各字詞在語意空間(Semantic Space)中的矩陣;  $S'^{-1}$  為語意空間之反矩陣(Inverse Matrix)。

圖 7.4 分隔線上方為建構原始矩陣  $X$  之資料流程, 主要執行動作為奇異值分解。在人工挑選的網路評論中, 將關鍵字在每篇評論中出現次數(Occurrence)之對應矩陣透過奇異值分解(SVD), 獲得預設的三個矩陣, 經過 dimension reduction 後得  $X \sim X' = T'S'D'^T$  作為預設的 SVD 矩陣。圖 7.4 下方為應用階段之資料流程。此機制將自動選擇領域詞庫中的詞彙, 透過文件前處理(Document Pre-processing), 將每篇擷取到的評論萃取出關鍵字之詞頻(Term Frequency), 並建立其在語意空間中之矩陣  $D_q$ , 如 F5.6 所示。

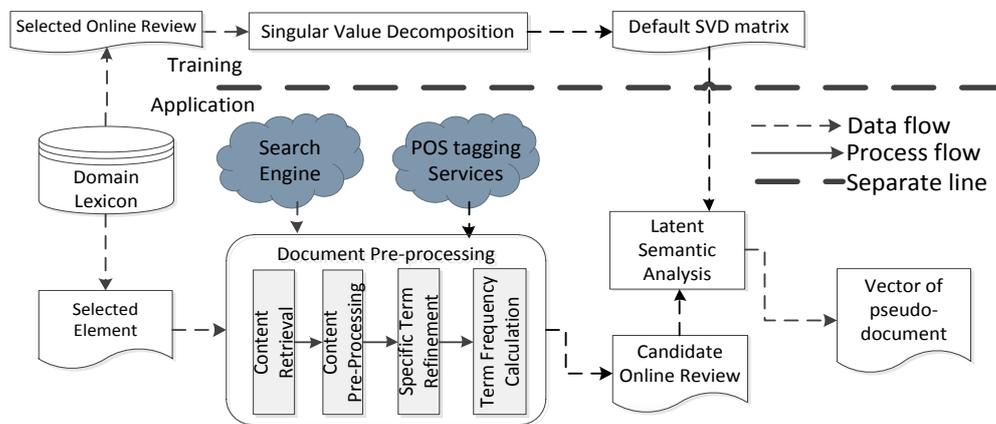


圖 7.4：Potential Customer Searching Process

### 7.4 潛在顧客識別

潛在顧客識別之程序(圖 7.5)說明如下：

- (1) Defining PCA：定義產品一般屬性(Product Common Attribute, PCA)，企業可依據產業知識，定義出產品共通的屬性。各屬性之概念描述應可用於描述企業之所有產品，例如品質(Quality)可適用於各種食品，外觀(View)可適用於所有產品等。
- (2) Mapping PCA to Consumption Value (CV)：消費價值理論的各項價值對於消費者而言皆有其影響層面，通常與產品之屬性(Attribute)相關，例如：衣服及珠寶可滿足消費者的情緒(Emotional)或社會(Social)價值，係因其高度外顯的性質所致。由此可知，「外觀屬性」與情緒及社會價值相關；又如 3C 產品，消費者購買前會考量如耐用度(Durability)、可靠度(Reliability)等之功能(Functional)價值，由此可知，「品質屬性」與功能價值相關。因此，若可藉由該產業之相關知識、資訊來定義產品一般屬性，建立與該理論關聯之假說，將有助於了解屬性(Attribute)與消費價值(Consumption Value)間之實際關聯。
- (3) Constructing weighting table：透過問卷題目之制訂、施測來檢驗上述對應關係之假說(Hypothesis)是否成立。將問卷之施測結果，利用結構方程模式(Structural Equation Model, SEM)進行檢驗，了解各路徑之因素負荷量(Factor Loading)及假說是否成立等。接著利用各路徑之因素負荷量產生消費價值與產品一般屬性對應之權重表(Weighting Table)，以供自動化環境參考。

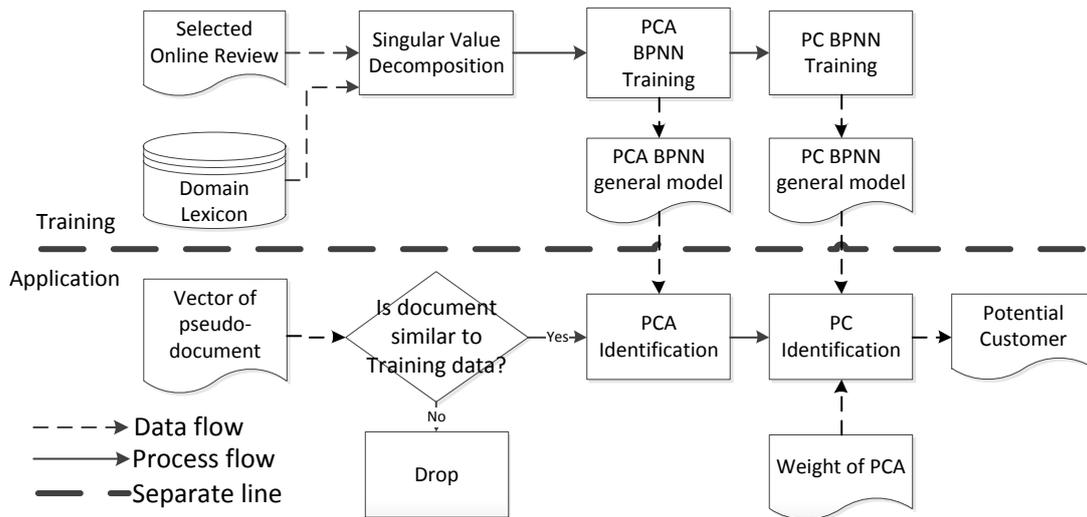


圖 7.5：潛在顧客識別程序

在圖 7.6 中，分隔線上方透過奇異值分解(SVD)，將  $D^T$  中各文件所對應之向量作為 BPN 模型左側之輸入( $X_1 \sim X_k$ )，利用給定之標記結果(即文件是否提及產品一般屬性之概念)訓練此 BPN 模型，收斂後各文件之輸出為其所對應之屬性矩陣。標記結果係由人工判斷各文章是否提及各屬性之概念，若文章中含有該概念則標記為 1，否則為 0。此 BPN 模型之輸入為維度  $k$  之向量，隱藏層  $n$  個神經元，輸出層為  $m$  個屬性(對應產品一般屬性)，神經元之間每條線均有其權重值，初始值隨機指定。

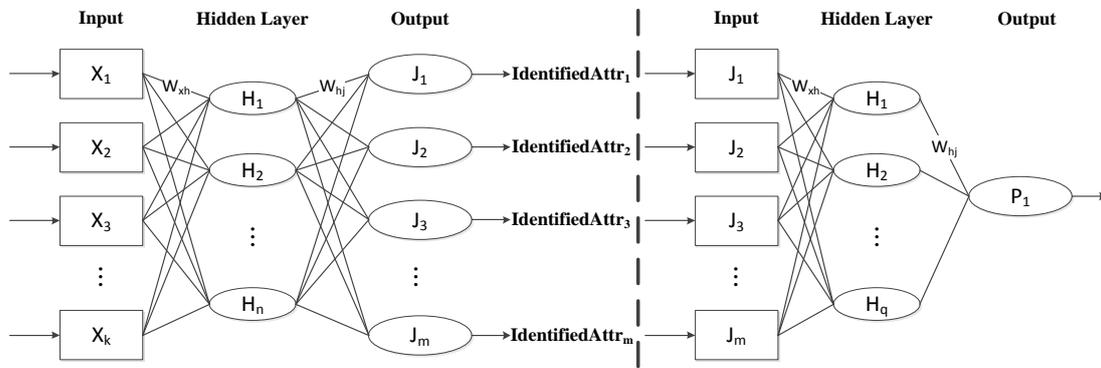


圖 7.6：BPN model of PCA/PC Identification

訓練完成之 BPN 模型將可利用收斂後之權重值進行預測，預測結果將輸出屬性矩陣，將此矩陣之預測值導入 weighting table，再次以 BPN 進行潛在顧客之識別，輸入值之轉換公式如 F7.7 所示、潛在顧客識別之 BPN 模型如圖 7.6 右側所示：

$$J_i = IdentifiedAttr_i \times (1 + Weight_i) \quad (F7.7)$$

其中， $J_i$  為此 BPN 模型第  $i$  個節點之輸入值； $IdentifiedAttr_i$  為前一階段 BPN 之輸出值、 $Weight_i$  為問卷調查第  $i$  個屬性對特定價值之路徑係數。

應用階段將訓練階段所得之  $D^T$  中各評論之向量進行 Cosine 相似度計算，相似度門檻值預設 0.5，可由使用者端調整。若上述步驟之判定結果為相似，則將作為第一階段 BPN 之輸入，識別該篇文章所含之概念。第一階段識別完成後，將已識別之屬性矩陣透過第二階段 BPN 進行識別，進一步判斷該篇文章中，作者是否為潛在顧客。其輸出結果與訓練階段不同，在此輸出結果為一介於 0~1 之數值，代表其為潛在顧客之可能性，越接近 1 則越有可能為潛在顧客，而輸出之門檻值由使用者決定，本研究預設為 0.5。

#### (八)結果與討論(含結論與建議)

本階段針對本報告歸納的三項主要研究產出，分別進行討論與建議，以利未來研究的持續改善的參考，分述如下：

- (1) 以社交運算改善虛擬社群交友安全之方法研究項目：本研究項目所調查之虛擬社群使用者網路行為，以廣義的網路使用者為研究對象，並未詳加區分研究對象是否以智慧型手機上網且為行動社群使用者；行動上網者因智慧型手機具有 GPS 定位功能，人們「搖一搖」手機即可知道陌生網友的人身位置，去除了網路的隔離特性，虛擬社群使用者所面臨的未知風險提高，其互動因子更形複雜。有關安全互動因素重要性分析，未來也許可以考慮採層級分析法(AHP)來進行。
- (2) 推論網路使用者人格研究項目：人格為網路使用者之內在屬性，其屬性值並不會出現在現有平台之個人公開資料中，需有方法藉由推論得知。本研究議題於本計畫第三年延伸出，目前共設計兩種方法分別為：特徵分類法及成分比對法，進行網路使用者推論工作。推論網路使用者人格之目的，是期能以此配合原先規劃改善虛擬社群交友安全之方法與虛擬世界實體人物辨識方法，提升網路使用者品質評估成效。本項目現階段僅完成設計方法流程與針對 Facebook 設計預設特徵集，分類器類型與方法有效性及推論精準度驗證等，仍有待後續研究的持續進行。
- (3) 潛在顧客搜索研究項目：在中文語意分析中，目前許多研究仍然針對正規格式之文件進行分析，對於主題不定、格式不定、用語不定的網路環境，因困難度頗高，目前仍較少研究者投入探討。本研究認為，欲於虛擬社群中自動地獲取有用的資訊，必得植基於厚實之語意分析能力。因此本項目將研究之重心著重於中文語意處理，透過提升中文處理之效率，不僅得以由虛擬社群中取得潛在顧客，更有助於後續延伸之應用。本研究項目在潛在顧客相關議題中，最先提出以自動化方

法於虛擬社群中搜索潛在顧客之概念。由於本研究項目之資料來源來自於網路，而使用者之個人資訊公開程度會依使用者的意願，使用習慣等而有所不同，這都將影響判定的正確性。另外，網路上充斥著各種網路特殊用語，雖然其本身通常不具複雜之意涵，但其存在形式廣泛，且隨著時事而變動。例如情緒符號、錯別字(Misspelt word)、諧音(homonyms)、與象形文字(pictograph)等。因此，未來有興趣的研究者可針對網路中這些有趣的文字現象進行進一步的探討。

綜觀本計畫的研究過程，適逢商務環境的變遷與政府法令更新，又受限於軟硬體資源及人力上的限制，致使一些原先規畫的研究項目在實際進行上有其困難度，因此只能在研究項目上做適當的調整，期望能夠將研究資源做最有效的應用。本計畫依調整之結果，依序執行之，本計畫結果與調整後之計劃工作項目大致相符。本年度的工作項目，主要完成了「虛擬社群互動安全評估的方法」及「推論網路使用者人格類型方法」的設計工作。並將「虛擬社群互動安全評估的方法」及「潛在顧客搜索機制」的研究產出，陸續整理撰寫成期刊論文，預計投稿國外期刊。

為期三年之計畫，藉由上述各項研究，得以順利完成，各階段之產出已陸續投稿至國內外期刊及研討會，列表如下：

● 研討會論文：

- (1) 陳宗義、劉彥辰、陳裕民(2012)，虛擬社群中潛在顧客搜索機制研發-以食品業為例，2012 資訊管理暨電子商務經營管理研討會。
- (2) 陳宗義、陳垂呈、陳俊男、劉彥辰(2012)，發掘網路拍賣欺騙喊價競標者：探勘技術之應用，2012 工業工程學會年會暨學術研討會。
- (3) 陳宗義、陳惠文，雲林縣國小學童使用臉書對其人際關係與學習動機之交互影響，第八屆國際健康資訊管理研討會，嘉義，南華大學，2013 年 6 月 14-15 日。
- (4) 陳宗義、陳松奇、陳裕民(2013)，虛擬世界實體人物搜尋方法發展，第九屆知識社群國際研討會。
- (5) Tsung-Yi Chen, Yan-Chen Liu, and Yuh-Min Chen, Research on Product Common Attribute model with Consumption Value Theory Applied in Food Industry, The IEEE International Conference on Industrial Engineering and Engineering Management, Bangkok, Thailand, December 10-13, 2013.
- (6) Tsung-Yi Chen, Hsiao-Chuan shen, A Network Behavior Quality Evaluation Method Based on Social Computing, 2014 International Conference on e-Commerce, e-Administration, e-Society, e-Education, and e-Technology, Nagoya, JAPAN, April 2-4, 2014.
- (7) Tsung-Yi Chen, Chien-Yun Huang, Analysis on the Differences in Internet User Experiences and Virtual Activities in Taiwan, International Conference on Engineering and Applied Science (2014 ICEAS), 22-24 July, 2014, Sappora, Japan.
- (8) Tsung-Yi Chen, Meng-Che Tsai, Yuh-Min Chen, Construction of an Interactive Behavioral and Feature Structure Model for Facebook, The IEEE International Conference on Industrial Engineering and Engineering Management, Kuala Lumpur, Malaysia, December 9-12, 2014.
- (9) A Product Positioning Analysis framework based on Maslow's Hierarchy of Needs Theory. (撰寫中)

● 期刊論文：

- (1) Tsung-Yi Chen, Yuh-Min Chen, Hui-Ying Cai, Design of Knowledge Trading Negotiation Model for Accelerating Acquirement of Knowledge, International Journal of Computer Integrated Manufacturing, Volume 25, Issue 12, 2012, pages 1085-1101. (SCI)
- (2) Tsung-Yi Chen, Yuh-Min Chen, Tai-Shiang Wang, Developing an ontology-based knowledge combination mechanism to customize complementary knowledge content, International Journal of Computer Integrated Manufacturing. (SCI) (Accepted).
- (3) Development of a Potential Customer searching Mechanism from Virtual Communities: Food Industry as an application, Knowledge-based Systems. (SCI) (審稿中).
- (4) 從網路犯罪觀點探討網路社群虛擬人際行為與網路互動安全決策因素，電子商務研究。(審稿中)

- (5) A 3-dimension Social Computing-based Evaluation Method for Improving Interactive Security in Virtual Network Worlds, Expert Systems with Applications. (SCI) (審稿中).
- (6) Development of a Product Positioning Analysis Mechanism based on Maslow's Hierarchy of Needs Theory, Electronic Commerce Research and Applications. (SCI) (撰寫中)

### 參考文獻

- Adalı S. Golbeck J. (2012). Predicting Personality with Social Behavior. International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 302-309.
- Adali S., Sisenda F., & Magdon-Ismael M. (2012). Actions Speak as Loud as Words: Predicting Relationships from Social Behavior Data. WWW '12 Proceedings of the 21st international conference on World Wide Web, 689-698.
- Alessandra, T., & O'Connor, M.J. (1998). The Platinum Rule: Discover the Four Basic Business Personalities and How They Can Lead You to Success.
- Bai, S.T., Zhu, T.S., & Cheng, L. (2012). Big-Five Personality Prediction Based on User Behaviors at Social Network Sites, eprint arXiv: 1204.4809.
- Berry, M. J., & Linoff, G. (1997). Data mining techniques: For marketing, sales, and customer support. John Wiley & Sons, Inc.
- Boyd, C.F. (1994). Different children different needs: The art of adjustable parenting. Multnomah.
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., Zanasi, A. (1998). International Business Machines Corporation (San Jose, California) & International Technical Support Organization (San Jose, California). Discovering data mining: from concept to implementation (Vol. 1). Upper Saddle River, NJ: Prentice Hall.
- Carver, C. (1999). Building a virtual community for a tele-learning environment, IEEE Communications Magazine, 37(3), 114-118.
- Chang, J. H. (2011). Mining weighted sequential patterns in a sequence database with a time-interval weight. Knowledge-Based Systems, 24(1), 1-9.
- Craig, D.L. & Zimring, C. (2000). Supporting Collaborative Design Groups as Design Communities, Design Studies, 21(2), 187-204.
- Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R. (1990). Indexing by latent semantic analysis, Journal of the American Society for Information Science, 41, 391-407.
- Erickson, T. (1997). Social interaction on the Net: virtual community as participatory genre, Proceedings of the Thirtieth Hawaii International Conference on System Sciences, 6, 13-21.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases, AI magazine, 17(3), 37.
- Golbeck J., Robles C. & Turner K. (2011). Predicting personality with social media. In Proceedings of the 29th ACM Conference on Human Factors in Computing Systems (CHI).
- Goring, C. B. (1913). The English convict: A statistical study. London: His majesty's Stationery Office.
- Gottfredson, M. G., & Hirschi, T. (1990). A general theory of crime. Palo Alto, CA: Stanford University Press.
- Hagel, J. & Armstrong, A. (1997). Net Gain: Expanding Markets through Virtual Communities, Harvard Business School Press.
- Hamberger, Y. A., & Ben-Artzi, F. (2000). The relationship between extraversion and neuroticism and the

- different uses of the Internet, *Computers in Human Behavior*, 16, 441-449.
- Han, J., & Kamber, M. (2001). *Data mining: Concepts and techniques*, China Machine Press, 8, 3-6.
- Hesse, B.W. (1995). Curb cuts in the virtual community: telework and persons with disabilities, *Proceedings of the Twenty-Eighth Hawaii International Conference on System Sciences*, 4, 418-425.
- Hirschi, T. (1969). *Causes of delinquency*. University of California Press.
- Cohen, M. L., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach, *American Sociological Review*, 44, 588-608.
- Ho, J., Schraefel, M.C. & Chignell, M. (2000). Towards an Evaluation Methodology for the Development of Research Oriented Virtual Communities, *Enabling Technologies: Infrastructure for Collaborative Enterprises*, *Proceedings of IEEE 9th International Workshops on, Enabling Technologies: Infrastructure for Collaborative Enterprises*, 112-117.
- Hong, C. F., Lin, M. H., Yang, H. F., & Huang, C. J. (2010). A novel chance model for building innovation diffusion scenario, *2010 IEEE International Conference on Systems Man and Cybernetics (SMC)*, 3845-3852.
- Jones, Q. & Rafaeli, S. (2000). Time to Split, Virtually: Discourse Architecture and Community Building as means to Creating Vibrant Virtual Metropolises, *International Journal of Electronic Commerce & Business Media*, 10(4), 214-223.
- Knapp, M.L. & Daly, J.A. (2002). *Handbook of Interpersonal Communication*, SAGE Publications.
- Kwai Fun IP, R., and Wagner, C. (2008). Weblogging: a study of social computing and its impact on organizations, *Decision Support Systems*, 45, 242-250.
- Lee, F.S.L., Vogel, D. & Limayem, M. (2002). Virtual Community Informatics: What We Know and What We Need to Know, *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 2863-2872.
- Liu, H. & Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. Boston: Kluwer Academic Publishers.
- Lu, Y., Zhao, L., and Wang, B. (2010). From virtual community members to C2C e-commerce buyers: trust in virtual communities and its effect on consumers' purchase intention, *Electronic Commerce Research and Applications*, 9(4), 346-360.
- MBAlib (2013). [http : //wiki.mbalib.com/zh-tw/DISC](http://wiki.mbalib.com/zh-tw/DISC).
- Messinger, P.R., Stroulia, E., Lyons, K., Bone, M., Niu, R.H., Smirnov, K., and Perelgut, S. (2009). Virtual worlds—past, present, and future: new directions in social computing, *Decision Support Systems*, 47, 204-228.
- Moulton, M.W. (1928). *Emotions of Normal People*, Kegan Paul Trench Trubner And Company.
- Ohsawa, Y., Benson, N. E., & Yachida, M. (1998). KeyGraph: Automatic indexing by co-occurrence graph based on building construction metaphor, *IEEE International Forum on Research and Technology Advances in Digital Libraries*, 12-18.
- Olson, D. L., & Shi, Y. (2007). *Introduction to business data mining*. New York: McGraw-Hill/Irwin.
- Ortigosa A., Carro R. M., & Quiroga J. I. (2014). Predicting user personality by mining social interactions in Facebook, *Journal of Computer and System Science*, 80, 57-71.
- Parameswaran, M, Whinston, A.B. (2007). Social computing: an overview, *Communications of the Association for Information Systems*, 19, 767-770.

- Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*, Addison Wesley.
- Romm, C. & Clarke, R.J. (1995). *Virtual Community Research Themes: A Preliminary Draft for a Comprehensive Model*, 6<sup>th</sup> Australasian Conference on Information Systems.
- Sakakibara, T., & Ohsawa, Y. (2007). Gradual-increase extraction of target baskets as preprocess for visualizing simplified scenario maps by KeyGraph, *Soft Computing*, 11(8), 783-790.
- Shekarpour, S., and Katebi, S.D. (2010). Modeling and Evaluation of Trust with an Extension in Semantic Web, *Web Semantics: Science, Services and Agents on the World Wide Web*, 8(1), 26-36.
- Srikant, R., & Agrawal, R. (1997). Mining generalized association rules, *Future Generation Computer Systems*, 13(2), 161-180.
- Uysal, A. K. & Gunal, S. (2012). A novel probabilistic feature selection method for text classification, *Knowledge-Based Systems*, 36, 226-235.
- Wang, F.Y, Carley, K.M, Zeng, D, Mao, W. (2007). Social Computing: From Social Informatics to Social Intelligence, *IEEE Intelligent Systems*, 22(2), 79-83.
- Wikipedia (2013). <http://zh.wikipedia.org/wiki/%E4%BA%BA%E6%A0%BC>.
- Wikipedia. (2014). <http://zh.wikipedia.org/wiki/Facebook>.
- Yand, J. H. & Honavar, V. (1998). Feature Subset Selection Using a Genetic Algorithm, *IEEE Intelligent Systems*, 13(2), 44-49.
- 洪叔民(2010), <<http://www.kmcc.org.tw/?content=casespec&id=1-10001-11>>。
- 張乃文(2009), Web2.0 網站平台管理之法制議題研析—以網路實名制與揭露使用者身份資料為中心, 《科技法律透》, 21(6), 42-60。
- 劉麗萍(2004),《以社會網路分析建構虛擬社群之人際信任指標》,東吳大學商學院企業管理學研究所, 碩士論文。
- 連征忠(2002),《虛擬社群成員間信任關係之探討》,大葉大學,事業經營研究所,碩士論文。
- 陳仲偉(2005),重思網路社群:網路主題樂園團體,《資訊社會研究》,8,271-286。
- 新北市政府警察局永和分局(2012),《防範犯罪宣導資料》, <http://www.yonghe.police.ntpc.gov.tw/web66/file/1350/upload/ct/ct63810w/wang.htm>。
- 法務部(2011),《法務部95-99電腦犯罪統計》, <http://www.moj.gov.tw/public/Attachment/152018152910.pdf>
- 廖有祿、李相臣(2003),電腦犯罪理論與實務,台北市:五南圖書出版,26-29。
- 馬傳鎮(2008),犯罪心理學新論,台北市:心理出版社。
- 施建彬(2007),誰在上網?網路行為及其相關因素探討研究,大葉大學共同教學中心,研究與動態, 16, 133-146。
- 創市際市場研究顧問(2010),網友花費在社群網站時間為電子郵件網站的9倍, [http://www.insightexplorer.com/news/news\\_11\\_23\\_10.html](http://www.insightexplorer.com/news/news_11_23_10.html)。
- 國家教育研究院雙語詞彙、學術名詞暨辭書資訊網(2014), <http://terms.naer.edu.tw/detail/1678987/?index=1>。