

植基於潛在語意索引之網路虛假評論辨識方法

陳隆昇

朝陽科技大學資訊
管理系

lschen@cyut.edu.tw

薛伶芸

朝陽科技大學資訊管
理系

ginspurple@gmail.com

摘要

隨著電子商務的發達，網路已經成為收集和分享消費者的個人意見、喜好、經驗與產品的一個很好的平台。基於普及的文本通信工具，客戶可以很容易地表達自己對購買的產品或服務的意見。一般來說，網路評論應該中立反應消費者對產品或服務的經驗。然而，有些廠商製造虛假評論以影響消費者行為，這會降低消費者的購買意願，並帶給企業極大的損害。儘管虛假評論已廣泛的存在網路評論中，但在文獻中此類相關研究卻很少。因此，本研究旨在透過特徵提取的技術，藉由使用 Global-LSI、Local-LSI 的重要功能，並定義及導入潛在虛假特徵，提升分類正確率，以精確找出虛假評論。本文實際在網路論壇上蒐集智慧型手機的評論進行實驗，以證實所提出方法的有效性。

關鍵詞：網路評論、操控偵測、支持向量機、維度縮減、潛在語意索引。

Abstract

With the proliferation of e-commerce, internet has become an excellent platform for gathering and sharing consumers' personal views on, preferences for, and experiences with products. With the popularity of text based communication tools, customers can easily express their opinions about purchased products or services. Generally speaking, the on-line reviews should be unbiased reflections of the consumers' experiences with the products or services. However, some comments are biased "manipulation", which might reduce consumers' purchase intentions and bring a great damage to enterprisers. This study aims to improve the performance for manipulation detection through reducing dimension space. We introduce feature extraction techniques, the important feature

is based on Global-LSI, Local-LSI is reduced by the dimension, to improve the detection accuracy further. A real case study of smart phone is used to illustrate the effectiveness of the proposed features.

Keywords: On-line comments, Manipulation detection, Support vector machines, Dimension reduction, Latent semantic indexing.

1. 前言

隨著 Web 2.0 的增長，已經快速的造成一股創新，逐漸改變購物者獲取資訊的方式，影響購物者購買產品或投資的決策 (Hu, Liu, and Sambamurthy, 2012)。研究顯示越來越多的消費者依靠網路獲取產品資訊和其他消費者的意見 (Utz and Kerkhof, 2012)。網路評論的產生者，即為同樣具有消費者身分的一般大眾，加上在網路上易於發表與擴散，這也是網路評論影響力超越其他傳統媒體的原因。同時，這些評論保存了消費者寶貴的意見與經驗，對企業來說也是重要的資訊之一。然而，在網路評論上提供的資訊是很容易被操控的，他們不得不尋求有信用的網路評論和網站資訊 (Pan and Chiou, 2011)。總括而言，網路社群口碑就是消費者的親身經驗分享，以心得或意見的形式存在於各大討論區、部落格與社群網站。

根據一份台灣資策會於 2013 年初發表的國內調查研究，網友購買商品會受到不同資訊來源影響，從 2012 年的 51.0% 增加到 57.0% 的網友會受「社群網站與部落客的分享」影響。由此可見，網路社群的影響力已默默地超越傳統媒體，成為左右消費者購買行為的主因。客戶認為這種用戶生成的評論網站的廣告更具有說服力。因此，了解消費者如何使用網路上這些被視為最有幫助的評論，並可以產生更高的產品判斷，是重要的課題 (Hu, Liu, and Zhang, 2008)。

Hu (2012) 發現越來越多的消費者依賴網路上的評論來做決策，利用檢測評論的寫作風格、情感、可讀性等操控屬性做實驗，發現網

路上約有 10% 的評論是屬於操控評論，企業發現商機的存在，便試圖去製造虛假評論(fake comments)，以影響消費者的購買行為。不幸的是，這些網路上“操弄”的虛假評論，可能會降低消費者的購買意願，並對企業帶來極大的損害。如何判別並過濾掉操弄的虛假評論，向客戶提供可靠的網站給客戶一個值得信賴的資訊是非常重要的。一個網站服務內容。

基於改善網路操弄評論分類績效之研究，本研究主要的目的將提出 11 個潛在虛假特徵屬性，並透過特徵提取方法提高分類性能的操作。用於檢測的基本模型是支持向量機(Support Vector Machines, SVM)。這項研究主要分為三個部分。在第一部分中，我們介紹了特徵提取的技術，藉由使用 Global-LSI、Local-LSI 等方法，試圖提高分類的性能。在第二部分中，我們導入了在文獻中整理的 11 個潛在的虛假特徵，並顯示這些外加的特徵可以顯著的提高檢測準確率。我們使用一個真實的案例智慧型手機來說明我們所提出的方法的效性。

2. 文獻探討

2.1 網路口碑

口碑(word-of mouth, WOM)最早是來自人際口耳相傳的結果。學者對口碑定義為，口碑傳播的內容並不限於品牌、產品或是服務，所有與產品相關的想法、對產品的意見、消費者親身經驗的分享等，都包含在內 (Blackwell et al., 2006)。如今，隨著網路普及運用，消費者口碑傳播行為逐漸從實體人際網絡轉移到虛擬空間，形成傳播範圍更大更遠的「網路口碑」。

隨著網際網路的發達，使得資訊傳播更加快速與方便，因為上網人口的增加，網路評論數量也越多，透過網路傳播媒介也增加網路口碑的影響力 (Wang et al., 2010)。因為網路口碑不同於廣告，是屬於非商業性資訊。而且此口碑是由消費者所創造的資訊，因為此資訊提供了消費者的直接經驗，進而對購買者的產品購買決策有很大的影響力 (Park et al., 2007)，並對企業帶來傷害。

網路評論係指消費者經由網路平台，發表或分享個人的直接經驗，並且提供他人自由的閱讀與分享這些評論或資訊 (Lin, 2012)。消費者認為從線上評論收集而來的資訊比從商業性來源取得的資訊，對產品會產生較高的興趣 (Bickar et al., 2001)，並可以做為未來購買產品時的參考依據。Hu (2008)發現，消費者在網

路上尋找產品評論時，信賴於其他使用者所發表生成的評論，消費者個人經驗與心得分享會影響購物者的購買決策 (Utz et al., 2012)。因此，購買者認為網路口碑，可信度高於廣告，並且是值得信賴的。然而，購物者面對大量的網路評論，如何判斷不是虛假評論，對於探討網路評論的可信度已成為熱門探討的議題之一 (Racherla et al., 2012)。

2.2 虛假評論

消費者藉由網路溝通，在網上尋找、分享、了解所需的產品資訊，並協助解決複雜的問題，為消費者提供正確的資訊和意見是必要的。而消費者取得產品訊息除了透過廣告外，亦可從線上評論 (Online Review) 獲取產品評價之相關資訊 (Lee, 2009)。

企業發現商機的存在，便試圖創造出有利於公司的評論去影響消費者的購買行為。行銷商和供應商都使用這一媒介，不斷監測消費者的網路評論，更雇用寫手假借客戶名義虛增網路評論，發布非真實的消息到留言板上，提高其產品的銷售量，此行為視為一種操控評論 (Review Manipulation) 的手法。相關研究指出網路上確實存在著操控評論，例如：(1) Hu (2012) 發現越來越多的消費者依賴網路上的評論來做決策，利用檢測評論的寫作風格、情感、可讀性等操控屬性做實驗，發現網路上約有 10% 的評論是屬於操控評論；(2) 公司不僅定期在網路論壇或聊天室發布自己產品的資訊及宣傳廣告，也積極的要求消費者在網路評論上給予產品正面的評價，擅自撤下負面評價；並透過親朋好友、同事甚至聘請專業人士在產品的銷售評論上一致給予好評的評論，影響購物者做購物的決策 (Mayzlin, 2006)。我們相信不道德的用戶發表假的網路評論，他們的評論有很高的評價或虛假的文字是一般消費者很難發現的 (Dellarocas, 2006)。以上整理的文獻都屬於操弄的文獻，操弄評論又分為操弄評論及虛假評論，操弄評論意指在評論上出現企業企圖誘導消費者購買的操弄文字，而虛假評論則是指在評論中不誠實的文字，本研究僅針對虛假評論做偵測，在實驗部份所提到的操弄評論是指定義的虛假部份。

雖然消費者了解虛假評論的存在，儘管對於虛假評論的懷疑，以及虛假評論的數量是未知的，消費者只能期望找到可信度高的網路論壇給予消費者誠實的評論。Evans (2010)認為，雖然每一個特徵可能不是唯一的虛假特徵，但是將它們組合起來就有可能成為虛假的評論。

本文整理了由 Ben Popken 在 2010 年提出的一篇文章敘說 30 個能找出操弄評論的方法，歸納在 13 個可能為辨識操弄特徵屬性中，以下表 1 所示。正因為這種組合的特徵將組成一個虛假的評論，所以應充分的定義虛假的特徵屬性。本文根據表 1 為基礎重新定義 11 個潛在虛假的特徵屬性如：價格 (Price)、推薦強度 (Strength of recommendation) (Lee et al., 2012)、作者頻率 (Frequency)、時間 (Time) (Hu et al., 2011b)、多餘資訊 (Redundant information)、情緒文字 (Emotionality word)、發布總篇數 (Occurrence times)、手機品牌 (Phone brands) (Huang et al., 2012)、資訊不對稱 (Asymmetric information)、專業知識 (Expertise)、曝光率 (Exposure) (Hu et al., 2008) 等。

表 1 操弄評論辨識方法整理表

特徵	方法
負面評論	2. 嚴防操弄的負面評論。
	10. 負面評論往往是一個具體的產品資訊來源。
	11. 負面評論可能是心理不平衡的消費者或試圖打壓另一家競爭對手的評論。
	19. 只有少數的評價是屬於非常正面的。
	24. 評論者回應負面的評論問題，通常都會說他們沒有這樣的問題，甚至是解釋是別人做錯了。
產品全名和型號	8. 評論者重複敘述著產品的全名和型號。這往往是一種企圖以廉價搜尋引擎結果的一種方法。
	18. 隨著特許的經營方式，所有評論者在幾天內已經互相評論了全國連鎖的位址。
	25. 評論竟然不敘述產品本身的任何東西。
	26. 評論中通常都會包括一個連結到他們網址的鏈結。但評論本身是無用的。
極端評等	17. 評論者在計劃生產的一顆星評論中張貼三種對於墮胎的看法。
	29. 完全避免 5 顆星和 1 顆星的級別，而趨向於 2~4 顆星。發現 2~4 顆星通常比較可靠。
行銷的語法	1. 行銷的語言，正常人不會使用行銷的口語撰寫評論。
明顯標記(字母大寫、旗標標註)	5. 即使他們不是假的評論者，對於全部寫大寫的人是種愚笨的行為，應該被忽略掉。

	30. 紅色的旗子：論述和推薦的任何變化。 (此指重點、特色或是吸引顧客注意的地方)
評論者的資訊	3. 評論者在這個網站沒有撰寫其他評論。
	21. 所有評論者都大約在同一個時間創建帳戶，通常在這時間左右發布評論。
	22. 該用戶名超過 3 個數字結尾。特別是其他的評論結尾時留下超過 3 個數字，通常這是自動系統撰寫評論時留下的記號。
重複的評論	4. 很多個評論都是完全相同的。
	27. 運用 google 的方式來搜尋較突兀或是特殊用詞的詞語，發現他們只是把千篇一律的評論，遍地開花的放在不同的評論網站。
特定品牌	7. 單一的評論者其唯一對一家製造廠商的產品進行評論。
	9. 評論者使用通過檢驗商標樣式的名字。
流行用語	20. 網路評論使用很多流行用語來描述其產品或服務。
產品發表前	23. 如果評論在實際產品的前幾個月或前幾周的製作之前發布。
粉絲爭辯	28. 我抱持懷疑的態度懷疑任何一個評論都有產品或公司的內定粉絲的戰爭。 例如，我恨它，當產品 A 有粉絲留下假的負面評論，和產品 B 的粉絲就會引起共憤。
語意評論	6. 雙向的評論可信度較高，當每一件產品只有單一的正面或負面評論，就可能是操弄的評論。
	12. 聲稱產品是第一且唯一。
	13. 評論給了折扣資訊，甚至直接告訴你去哪裡購買該產品。
	14. 評論中進行長篇大論解釋為什麼該產品比其他相同但不同品牌的產品優秀。
	15. 評論看起來像是廢話，甚至文不對題。
	16. 為了得到報酬而撰寫的簡短評論。 例如：我喜歡他們的食物。

出處：

<http://consumerist.com/2010/04/14/how-you-spot-fake-online-reviews/>

2.3 特徵提取

特徵提取 (Feature extraction) 即是一種轉換的方法，可以產生一個新的特徵集合，主要是將原始特徵空間轉換或組合成不同於原始的新特徵空間 (Yang et al., 2012)。王金鳳 (2004) 學者使用特徵聚合理論 (FA)、潛在語意檢索 (LSI) 進行文本分類實驗，資料結果顯示 LSI 的分類方法比 FA 方法來的好。本研究採用的特徵提取方法為 LSI (Latent Semantic Indexing) 方法。

LSI 常被使用作為資料探勘的預先處理步驟，主要是具有計算上的優勢，因此在資訊檢索中也成為廣泛的使用工具 (曾韋榮, 2006)。此外 LSI 也能改善樣本資料稀疏的問題 (杜逸普, 2008)。

LSI 是建構在向量空間模型 (Vector Space Model) 的基礎上，並以奇異值分解 (Singular Value Decomposition, SVD) 的方法 (吳忻萍, 1997; 曾韋榮, 2006)，將原詞彙-文章向量矩陣，分解投射到維度較小的向量空間，將相關的詞彙及文章聚集在一起，以解決資訊檢索中，隱含語意的問題 (杜逸普, 2008)。

本研究將 LSI 視為一種特徵提取的工具，希望能藉由 LSI 解決在資訊檢索上的問題，以及向量空間維度縮減的特性，進行維度縮減。我們分別採取 Global-LSI 及 Local-LSI 兩種方法進行實驗。Global-LSI 做 SVD 矩陣分解是全部矩陣去分解，假設有第一類資料 40 筆，第二類資料 60 筆，共 100 筆資料；Global-LSI 是將 100 筆資料做 SVD 矩陣分解；而 Local-LSI 的作法是針對不同類別進行分解。意思是 Local-LSI 將 40 筆資料及 60 筆資料的矩陣分別做 SVD 矩陣分解。與 Global-LSI 只對整體資料文件矩陣進行一次 SVD 相比，Local-LSI 要進行 n 次 SVD，增加了計算次數，每一類的資料文件矩陣同整體資料文件矩陣相比要小的多，因此提高了計算的速度。

2.4 支持向量機

支持向量機 (Support Vector Machines, SVM) 是由 Vladimir Vapnik 在 1979 年提出的一種以統計學習理論 (Statistical Learning Theory) 為基礎的機器學習演算法，屬於監督式學習 (Vapnik, 1982)。由於 SVM 具有完整的理論架構，又有發展成熟的工具，且適用於小樣本、高維模式以及非線性的問題，只需要透過幾個關鍵的 support vector 即可進行分類，因此廣泛的被應用 (Cristianini et al., 2000)。對於支持向量機而言，原始的資料群假設為線性可分的情

況就可以使用超平面進行分類的動作，但是一般來說資料大多都是屬於線性不可分的情形，如果資料為線性不可分如圖 1(左)圖所示，則需要經過核心函數 (Kernel Function) 的映射改變資料型態，將輸入資料由原本較為低維的空間透過核心函數的映射轉到更高維的空間中做處理如圖 1(右)圖所示，進而解決線性不可分問題。

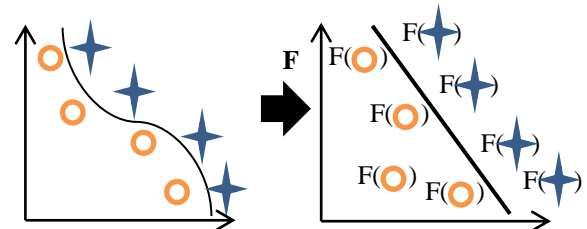


圖 1 (左)圖資料為線性不可分情形 (右)圖 (左)透過核心函數映射之後呈現線性可分的情形

(資料來源：修改自陳佳威, 2011)

在 SVM 執行的過程中會先由訓練資料透過決策函數 $D(x)$ 去計算出一條線，把資料切割成兩類的一個超平面 (hyperplane)，距離這條超平面附近的幾筆資料，只要帶入 $D(x)$ 中所得到的值為 +1 或 -1 的，就是所謂的 support vector，即為圖 2 中的兩條虛線。由於經過計算可能會得到不只一條 hyperplane 與 support vector 將資料一分為二，因此要找到兩條虛線之間的邊界 (margin) 越大的越可以有效率的切割兩個集合，因此，SVM 會找到 margin 為最大的 hyperplane，也就是最理想的超平面，可將資料做最明確的分類。

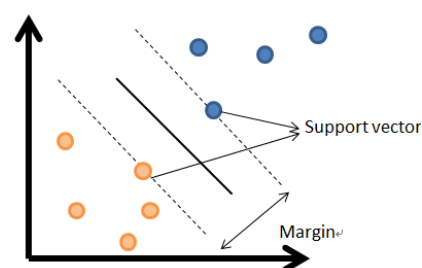


圖 2 SVM 示意圖

(資料來源：修改自 Unler et al., 2010)

3. 研究方法

本章節將介紹本研究實驗之流程。首先，我們使用 TF-IDF 權重，刪除低詞彙頻率後做為後選屬性集合進行實驗一。接著將實驗一資料導入 11 個潛在虛假特徵屬性進行實驗二。針對所設定的兩項實驗分別進行 Global-LSI 及

Local-LSI，兩種維度縮減的方法進行實驗。依等比例方式(50%、25%、20%、15%、10%、5%、4%、3%、2%、1%)，來縮減二類別特徵集合的維度，並分別選取重要度較高的特徵，以做為代表性特徵。

3.1 實驗流程

本研究詳細的步驟說明如下。以下為本研究操控評論辨識的實驗流程圖如圖 3。

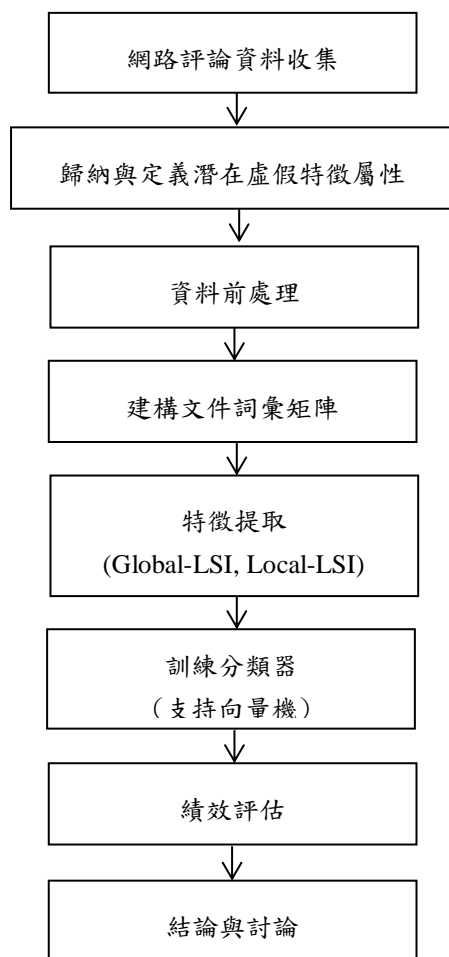


圖 3 虛假評論辨識流程圖

步驟一：網路評論資料收集

我們從論壇中分別收集智慧型手機的評論，根據資料集所探討的主題產品、設備與相關資訊，潛在虛假特徵的屬性也有所不同。

步驟二：歸納與定義潛在虛假特徵屬性

首先，根據文獻本研究歸納出 11 個潛在虛假特徵的屬性值。並根據表 1 的方法定義類別，例如：從負面評論可能為心理不平衡的消費者或試圖打壓另一家競爭對手評論中定義出情緒文字 (Emotionaluty word)；曝光率 (Exposure) 往往是一種企圖以廉價搜尋引擎結果得一種方法，評論者不斷重複著產品的全名及型號；時間 (Time)，評論收集時間是從 2011

年 11 月到 2013 年 9 月，故將時間分為前段小於 1 年、中段等於 1 年與後段大於 2 年；專業知識 (Expertise) 以資料集主題的智慧型手機找出 Android、Phone、16G、32G、microSD、ios... 等智慧型手機相關特徵。發布總篇數 (Occurrence times)，企業為了提高產品銷售量，雇用寫手達到曝光率，短時間內同作者在一個論壇發布多個一模一樣的評論吸引讀者等等。接著，根據所找出的潛在虛假特徵進行建構詞彙文件矩陣 (TDM) 並使用分類器進行實驗。

步驟三：資料前處理

首先，收集所需之原始資料，以 uni-gram 做為一個特徵單位，從所收集的文件資料擷取關鍵詞特徵。在資料的前處理部分，本研究使用 stop words 表來刪除文件中的冗詞贅字，與低詞頻的詞彙，接著刪除文件第一列資料。主要目的是在淨化資料與轉換實驗分析所需的格式。再使用分類器協助進行操弄評論的篩選，使用五折交叉驗證法以證實所提出方法之有效性。

步驟四：建構文件詞彙矩陣

本實驗透過 TF-IDF (Term Frequency-Inverse Document Frequency)，如公式 (3.1) 所示的權重來建構詞彙文件矩陣 (Term-Document Matrix, TDM)，如表 2 所示。

$$TF-IDF = TF(t_i, d_j) \times \log\left(\frac{N}{N(t_i)}\right) \quad (3.1)$$

在公式 3.1 中， t_i 為第 i 個特徵； d_j 為第 j 篇文件； N 為所有文件的總數； $N(t_i)$ 表示有出現 t_i 特徵的文件總數。

表 2 詞彙文件矩陣範例

	特徵 X ₁	特徵 X ₂	特徵 X ₃	...	特徵 X _n	類別 Y
文件 1	0	0	1.11 69		1.19 93	1
文件 2	2.39 86	2.01 22	0		0	1
文件 3	1.19 93	0	2.23 39		0	-1
...						
文件 n	2.04 96	2.39 86	0		0	-1

步驟五：特徵提取

LSI 由 Deerwester et al., (1990)所提出，藉由統計方法提取出文章的潛在語意架構，使得資訊檢索具有語意的關聯。而 LSI 主要是使用奇異值分解 (Singular value decomposition, SVD)將建構好的詞彙文件矩陣投射到維度較小的向量空間中，以向量空間作為分析模型。

1.Global-LSI

步驟 1 原始矩陣 A 分解：利用 SVD 將所蒐集的原始詞彙-文章向量矩陣 A 分解。

步驟 2 縮減 A_k 矩陣：取 k 維度，縮減為新的 A_k 矩陣。

步驟 3 資料分類預測：以縮減維度後的 M_{k-1} 矩陣，取代原始矩陣 A 作為分類預測之輸入資料，進行分類預測。

$$M_k = U_k \times S_k$$

其中， M_{k-1} 為維度縮減條件屬性維度的資料矩陣、 U_{k-1} 為 k-1 維度資料矩陣、 S_{k-1} 為 k-1 維度對角矩陣，且 $k-1 \leq k \leq n$ 進行實驗分類。

步驟 4 獲得分類預測數據並結束。

2.Local-LSI

步驟 1 群聚矩陣分解：Local-LSI 的作法有別於 Global-LSI 最大的差別，便是步驟 1。Global-LSI 做 SVD 矩陣分解是全部矩陣去分解，假設有第一類資料 40 筆，第二類資料 60 筆，共 100 筆資料；Global-LSI 是將 100 筆資料做 SVD 矩陣分解；而 Local-LSI 的作法是針對不同類別進行分解。意思是 Local-LSI 將 40 筆資料及 60 筆資料的矩陣分別做 SVD 矩陣分解。步驟 2 縮減為 A_k 矩陣：為了區分 Global-LSI 的 A_k 縮減矩陣，我們將 Local-LSI 縮減為 k 維度的矩陣命名為 A_{k2} 縮減矩陣。Global-LSI 的 A_k 矩陣是將原始矩陣分解後，取 k 維度即可得；而 Local-LSI 的 A_{k2} 矩陣則是將各群聚分解的矩陣，按照原始矩陣的資料順序排列回去。承步驟 1 的例子，將 40 筆資料縮減為 $A1$ 矩陣，將 60 筆資料縮減為 $A2$ 矩陣，最後將 $A1$ 及 $A2$ 按照 A 矩陣的順序排列回去，成為 A_{k2} 縮減矩陣。

步驟 3 資料分類預測：首先先將資料分回原本的第一類別資料 P 及第二類別資料 N，接著個別進行縮減維度產生 P_{k-1} 矩陣及 N_{k-1} 矩陣，再將 P_{k-1} 矩陣及 N_{k-1} 矩陣合併為 M_{k-12} 矩陣取代原始矩陣 A 作為分類預測之輸入資料，進行分類預測。

$$P_{k-1} = U_{k-1} \times S_{k-1}$$

$$N_{k-1} = U_{k-1} \times S_{k-1}$$

$$M_{k-12} = P_{k-1} + N_{k-1}$$

其中， M_{k-12} 為維度縮減條件屬性維度的資料

矩陣、 U_{k-1} 為 k-1 維度資料矩陣、 S_{k-1} 為 k-1 維度對角矩陣，且 $k-1 \leq k \leq n$ 進行實驗分類。

步驟 4 獲得分類預測數據並結束。

步驟六：訓練分類模型與績效評估

分割訓練與測試資料，使用 SVM 建構分類模型以進行分類，並執行效能評估 (Positive Accuracy, PA、Negative Accuracy, NA、Overall Accuracy, OA、G-mean, GM、Precision & Recall, F1)來驗證實驗結果。

步驟七：結論與討論

經分類器得到實驗數據後，再以平均值與標準差進行結果分析。

4. 實驗與結果

4.1 資料收集

本研究以手機品牌網路評論作為研究案例資料，使用 Amazon.com (<http://www.amazon.com/>) 主題為 Nokia Lumia 920、HTC One、Apple iPhone、Google Nexus 4 與 Samsung Galaxy S4 之消費者回覆評論，評論收集時間是從 2011 年 11 月到 2013 年 9 月。擷取的資料量與型態分佈如表 3 所示。

接續使用 QDA Miner 資料探勘軟體進行資料前處理，並建構實驗所需之詞彙文件矩陣。分類器在 SVM 的部分使用 Chang and Lin (2001)所開發的 LIBSVM 分類器。在 SVM 分類器中，本研究利用 libsvm 中預設的 grid search 功能，選取最佳 cost 與 gamma 參數用來訓練分類模型。

表 3 資料來源

資料主題	資料來源	文件數量
Nokia Lumia 920 : 389 筆 HTC One : 75 筆 Apple iPhone 5 : 320 筆 Google Nexus 4 : 499 筆 Samsung Galaxy S4 : 157 筆	Amazon.com	1440 筆

4.2 資料前處理

本研究依據文獻整理出以下 11 個潛在虛假屬性特徵，並給予定義，如表 4 所示。進行資料的前處理與區分出非操弄評論與操弄評論 (表 5 資料)

文件數量	類別分佈	資料來源
1440 筆	非操弄評論 : 1040 筆 操弄評論 : 400 筆	http://www.amazon.com/

接續以 uni-gram 為特徵單位進行詞彙擷取，並刪除無效評論與低詞頻詞彙。此外，本研究使用五折交叉驗證法 (5-fold cross-validation) 來驗證所提出方法之有效性。

該方法是將資料集分割成五等份，其中一份為測試用資料集，其他四份為訓練用資料集。因此，每項資料集將進行五次實驗。

表 4 潛在虛假屬性特徵值

編號	屬性	定義	文獻	屬性值定義
1	價格(Price)	評論中透露出影響讀者與價格有關的資訊	Lee et al.(2012)	1:有提到價格 0:無
2	作者頻率(Frequency)	同作者在同款手機論壇上發布的篇數	Hu et al.(2011b)	1:頻率 ≥ 2 次 0:頻率=2次
3	多餘資訊(Redundant information)	重複發布的評論或文章	Huang et al.(2012)	1:多餘資訊 ≥ 2 次 0:多餘資訊=2次
4	時間(Time)	將收取資料的時間從中間分為前段與後段	Hu et al.(2011b)	1: < 2年 0: =3年 -1: >5年
5	資訊不對稱(Asymmetric information)	消費者傾向於有較多的評論的產品，認為有較多的評論是更值得信賴的	Hu et al.(2008)	1:大多數的資訊 0:無
6	專業知識(Expertise)	評論出現專業的知識，吸引讀者目光	Hu et al.(2008)	1:有提到專業知識 0:無
7	情緒文字(Emotionality word)	情緒性文字影響讀者情緒並引導其意向	Huang et al.(2012)	1:有情緒文字 0:無
8	曝光率(Exposure)	消費者會忽略曝光率較低的評論或評論較少的產品	Hu et al.(2008)	1:高 0:無
9	發布總篇數(Occurrence times)	同一個帳號作者發布評論的數量	Huang et al.(2012)	連續值
10	手機品牌(Phone brands)	在評論中出現不同產牌的評論會正面影響讀者	Huang et al.(2012)	1:Apple 2:Samsung 3:HTC 4:Sony Ericsson 5:Nokia
11	推薦力強度(Strength of)	用強烈正向的文字加深使用者對產品的興趣	Lee et al.(2012)	1:強 0:中

	recommendation)			-1:弱
--	-----------------	--	--	------

表 5 資料

文件數量	類別分佈	資料來源
1440 筆	非操弄評論：1040 筆 操弄評論：400 筆	http://www.amazon.com/

4.3 資料收集

本研究存在著資料不平衡問題，為了有效的評估分類的效能，本研究分別使用準確率 (Accuracy, AC)、F1 和 Geometric mean (GM) 指標來評估分類效能並證明實驗的有效性。評估指標將藉由一個二元評估指標矩陣來計算，如表 6。

表 6 二元評估指標矩陣

Actual (實際)	Predicted (預測)	
	Positive (操弄)	Negative (非操弄)
Positive (操弄)	TP (the number of True Positive)	FP (the number of False Positive)
Negative (非操弄)	FN (the number of False Negative)	TN (the number of True Negative)

在表中，TP、FP、TN、FN 分別定義如下：

- 1).TP 表示操弄評論被正確分類成操弄評論的數量。
- 2).FP 表示操弄評論被錯誤分類成非操弄評論的數量。
- 3).TN 表示非操弄評論被正確分類成非操弄評論的數量。
- 4).FN 表示非操弄評論被錯誤分類成操弄評論的數量。

接著使用傳統評估指標均會使用的總準確率 (Overall Accuracy, OA)，來評估分類效能，其 OA 指標如公式(4.1)所示：

$$Overall Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4.1)$$

另外，根據 Zhang (2008) 學者指出 F1 指標方式結合了 Recall 和 Precision，可以有效評估分類的效能，F1 指標的計算方式如公式(4.4)所示：

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$Precision = \frac{TP}{TP + FP} \quad (4.3)$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4.4)$$

Geometric mean (GM) 指標是由 Kubat et al. (1997) 所提出的評估指標方法，主要是計算正向類別文件的準確率 (Positive Accuracy, PA) 和負向類別文件的準確率 (Negative Accuracy, NA) 之間的幾何平均數來評估分類效能。GM 指標的計算方式如公式(4.7)所示：

$$PA = \frac{TP}{TP + FN} \quad (4.5)$$

$$NA = \frac{TN}{FP + TN} \quad (4.6)$$

$$GM = \sqrt{PA \times NA} \quad (4.7)$$

GM 指標主要是求兩個類別準確率的最佳化，當 PA 和 NA 都很高，表示 GM 指標也是高的。

4.4 實驗結果與分析

實驗經支持向量機分類器得出實驗結果數據，同時經由五折交叉驗證法實驗五次得出實驗數據。本研究分別進行三項實驗：

實驗一：使用原始資料來建構詞彙文件矩陣

使用 TF-IDF 權重，刪除低詞彙頻率後做為候選屬性集合，最後建構詞彙文件矩陣。

實驗二：使用原始資料導入 11 個潛在虛假特徵來建構詞彙文件矩陣

將實驗一導入所定義出的 11 個潛在虛假特徵屬性值，建構出詞彙文件矩陣。

針對所設定的兩項實驗分別進行 Global-LSI、Local-LSI，兩種維度縮減的方法。依等比例方式(50%、25%、20%、15%、10%、5%、4%、3%、2%、1%)，來縮減二類別特徵集合的維度，並分別選取重要度較高的特徵，以做為為代表性特徵。

表 7 實驗一、使用原始資料

特徵維度	PA	NA	OA	GM	F1	Time (secs)
原始資料	13.2 5%	96.9 2%	35.3 9%	73.6 8%	21.6 0%	23.9
方法一- Global-LSI						
50%	11.5 0%	97.3 1%	33.3 7%	73.4 7%	19.4 0%	19.15
25%	13.0 0%	96.8 3%	35.1 4%	73.5 4%	21.2 2%	17

20%	11.2 5%	97.8 8%	32.5 6%	73.8 2%	19.0 1%	15.93
15%	13.0 0%	97.6 9%	35.4 1%	74.1 7%	21.7 7%	16.08
10%	17.5 0%	94.9 0%	40.6 4%	73.4 0%	26.7 4%	13.45
5%	13.7 5%	98.2 7%	36.6 1%	74.7 9%	23.1 7%	12.78
4%	12.7 5%	97.7 9%	34.7 2%	74.1 7%	21.2 2%	16.95
3%	12.0 0%	97.0 2%	33.7 0%	73.4 0%	19.8 4%	12.25
2%	11.0 0%	96.5 4%	31.8 4%	72.7 8%	17.8 9%	12.82
1%	12.0 0%	96.7 3%	33.8 1%	73.1 9%	19.8 0%	13.21
方法二-Local-LSI						
20%	73.7 5%	97.0 2%	84.3 8%	90.5 6%	80.9 7%	16.2
15%	75.5 0%	98.2 7%	85.9 6%	91.9 4%	83.5 9%	15.9
10%	79.2 5%	98.1 7%	88.1 0%	92.9 2%	85.9 9%	13.25
5%	84.7 5%	96.8 3%	90.4 9%	93.4 7%	87.6 9%	11.9
4%	86.2 5%	96.7 3%	91.2 9%	93.8 2%	88.4 9%	9.85
3%	85.0 0%	97.1 2%	90.7 9%	93.7 5%	88.1 9%	8.7
2%	82.2 5%	96.4 4%	89.0 0%	92.5 0%	85.7 7%	8.25
1%	80.0 0%	96.4 4%	87.8 0%	91.8 8%	84.4 9%	6.6

表 8 實驗二、使用原始資料導入 11 個潛在虛假特徵

特徵維度	PA	NA	OA	GM	F1	Time (secs)
原始資料	84.7 5%	95.6 7%	90.0 4%	92.6 4%	86.4 8%	35.12
方法一-Global-LSI						
50%	84.5 0%	95.2 9%	89.7 2%	92.2 9%	85.9 1%	29.5
25%	82.5 0%	95.6 7%	88.8 3%	92.0 1%	85.1 9%	27.2
20%	83.0 0%	96.0 6%	89.2 6%	92.4 3%	85.8 6%	25.1
15%	81.2 5%	96.2 5%	88.3 8%	92.0 8%	85.0 3%	25

10%	85.0 0%	95.1 9%	89.9 4%	92.3 6%	86.0 8%	23.6
5%	87.0 0%	96.2 5%	91.4 8%	93.6 8%	88.4 1%	16
4%	88.7 5%	96.2 5%	92.4 0%	94.1 7%	89.4 1%	17.45
3%	88.7 5%	95.9 6%	92.2 7%	93.9 6%	89.0 7%	20.15
2%	89.0 0%	96.2 5%	92.5 3%	94.2 4%	89.5 5%	20.2
1%	88.2 5%	96.1 5%	92.0 7%	93.9 6%	88.9 7%	4.45
方法二-Local-LSI						
20%	88.5 0%	98.5 6%	93.3 7%	95.7 6%	92.0 5%	23.4
15%	89.7 5%	97.6 0%	93.5 7%	95.4 2%	91.5 7%	24.9
10%	89.0 0%	97.5 0%	93.1 2%	95.1 4%	91.0 2%	23.5
5%	90.7 5%	97.0 2%	93.8 1%	95.2 8%	91.4 4%	22
4%	91.7 5%	97.4 0%	94.5 2%	95.8 3%	92.4 5%	21.6
3%	91.0 0%	98.5 6%	94.7 0%	96.4 6%	93.4 4%	11.85
2%	91.5 0%	98.2 7%	94.8 1%	96.3 9%	93.3 6%	8.85
1%	91.5 0%	98.4 6%	94.9 0%	96.5 3%	93.5 9%	4.05

從 7 表看 SVM 分類器實驗結果，以 GM 為主要評估指標。首先，發現原始資料 OA 分數為 35.39%，GM 為 73.68%，PA 為 13.25%，因此，從此結果可得知在正常的評論之下，只能判斷出些微的虛假評論，接著，我們進行兩個方法維度縮減的實驗，Global-LSI 方法維度縮減後最好的維度在 5%，提昇到 74.79%。而 Local-LSI 方法維度縮減後最好的維度在 4%，提昇到 93.82%。從以上分數得知，維度縮減後，只能找到少部分的虛假評論，所以我們導入了 11 個潛在的虛假特徵屬性，進行實驗二。

再從表 8 看 SVM 分類器實驗結果，一樣以 GM 為主要評估指標，首先，原始資料從 73.68% 提升到 92.64%，而 Global-LSI 方法從 74.79% 提昇到 93.68%；Local-LSI 方法從 93.82% 提升到 95.83%。實驗資料顯示，Local-LSI 方法擁有最佳的分類效果。

5. 結論

隨著網際網路的快速發展，網路平台具有公開性、即時性和共享的原則，成為人們表達意見、感受情緒、分享經驗的趨勢以及改變消費者過去購買商品接收新資訊的方式。然而網路的評論並不是完完全全是真實由使用者所發表的評論。因此，本研究主要的目的在於，找出能檢測虛假評論的重要操弄特徵屬性值，藉由這些特徵在未來可以做更多虛假評論的檢測，並找出更多的虛假評論來幫忙消費者做評論真偽的把關。

本研究透過文獻尋找出虛假相關的特徵屬性值，並對操弄評論進行檢測。以手機品牌網路評論作為研究案例資料，使用 Amazon.com 網站收集手機產品評論資料，主題為 Apple、Samsung、HTC、Sony Ericsson 與 Nokia 之消費者回復評論。根據實驗結果整理出以下結論：

1. 使用 SVM 分類器進行檢測虛假評論，結果顯示特徵提取方式以 Local-LSI 方法最佳，雖然提高分類率，但只能判斷出部分的虛假評論，對於虛假評論的效用不高。
2. 使用特徵提取後的特徵導入 11 個潛在虛假特徵屬性，擁有不錯的分類績效。這意味著本研究所提出的 11 個潛在的虛假特徵屬性，可以有效的幫助機器學習檢測分類出操弄評論。

綜合以上的結論，可以發現面對不同網站不同主題的單一性或混合性資料集，虛假特徵屬性具有一定程度的重要性，更能提高分類的準確度。

致謝

本文受到國科會專案計畫部分資助(契約編號 NSC 101-2628-E-324-004-MY3)，作者在此表達感謝之意。

參考文獻

- [1] 杜逸普 (2008)，部落格探勘—以網路電話產品為例，碩士論文，朝陽科技大學資訊管理系，霧峰，台中。
- [2] 周嵩能 (2010)，以微網誌語料進行情緒辨識之研究，碩士論文，國立台南大學資訊工程學系，台南。
- [3] 曾韋榮 (2006)，結合潛在語意檢索及資訊粒化於資料探勘，碩士論文，國立台北科技

大學商業自動化與管理研究所，台北。

- [4] 鄧振樹 (2013)，建置企業產品知識主題地圖—以 N 公司為例，碩士論文，國立成功大學工業與資訊管理學系，台南。
- [5] 周士堯 (2013)，應用文件探勘技術進行立法文本自動化分析，碩士論文，國立清華大學服務科學研究所，新竹。
- [6] Asdrúbal L. C., Jair C., Lourdes L. G., Farid G. L. Fisher's decision tree, *Expert Systems with Applications*, vol.40, pp.6283-6291, 2013.
- [7] Cortes, C. and Vapnik, V. N., "Support-Vector Networks", *Machine Learning*, 20, 1995.
- [8] Chrysostomou, K., Chen, S. Y., and Liu, X., Identifying user preferences with wrapper-based decision trees, *Expert Systems with Applications*, vol.38, pp.3294-3303, 2011.
- [9] Decker, R., Trusov, M., Estimating aggregate consumer preferences from online product reviews, *Intern.J. of Research in Marketing*, vol.27, pp.293-307, 2010.
- [10] Dellarocas, C., Strategic manipulation of internet opinion fo-rums: implications for consumers and firms, *Management Science* vol.52 (10) pp.1577-1593, 2006.
- [11] Freedman, L., Merchant and customer perspectives on customer reviews and user-generated content, 2008.
- [12] Hu, N., Liu, L., and Sambamurthy, V., Fraud detection in online consumer reviews, *Decision Support Systems*, vol.50, pp.614-626, 2012.