

部落客內隱關係之社會網路分析研究

黃夙賢

明新科技大學資訊管理系

shuang@must.edu.tw

林炳憲

明新科技大學資訊管理系

C96050002@std.must.edu.tw

摘要—Web2.0 概念提出，使得社群網站興起，例如部落格就為重要的網路服務之一。本研究目的為二，一為探討單一部落客是否會擁有多個部落格，二為探討部落客間的外顯、內隱社會網路於合併前後是否會產生不同。本研究針對無名與痞客邦部落格使用帳號、站名、文章分類與文章來分析部落格間的相似度來找尋部落客。接著再透過帳號與關鍵字查尋找出其部落客稱謂，並以此找到部落客間的內隱關係，最後將部落客間的關係資訊轉繪成社會網路關係圖，再使用圖形分析部落客間的社會網路。結果發現部落客社會網路在加入內隱關係後，原社會網路關係發生改變，使得孤立部落客因此與群體產生關聯。本研究提出線上社會網路分析的內隱關係分析模式，提供給從事部落客語意分析更多元的內隱關係。

關鍵詞—社會網路分析、部落格、內隱關係

Abstract—Web 2.0 arises the trend of community websites, and makes the weblogs to gain popularity and to be one of the most popular services. Purposes of this paper have two aspects. First, detect the blogger who owns multiple blogs. Second, study the blogger social network after join the explicit and implicit relationships. This paper evaluates the weblogs in Wretch and Pixnet, and finds out the blogger with multiple weblogs by account name, name of blog, article categories and articles similarity. Hereafter, discover the implicit relationships among bloggers by using blogger's appellation retrieved from account and keyword. The relationships are finally transformed to obtain sociograms among bloggers. The experiment result showed that the sociogram are changed after adding implicit relationships. The change joins some

isolated bloggers into related communities. This paper proposes analysis model to derive various implicit relationship for those who are willing research in online social network analysis.

Keywords—Social network analysis, Weblog, Implicit Relationship

一、簡介

部落格(Weblog，又可縮寫成 Blog)，亦被稱為網誌、博客，首先由 Jorn Barger 所提出。Jorn Barger 將個人認為有趣的資訊轉貼於個人網頁中，並將最新的資訊放置於整個網頁的最上方，最舊的資訊則放置於最下方，方便於使用者瀏覽相關訊息[5]。這種風格的網頁因此掀起了 Web2.0 的革命，使得社群網站興起，因此部落格就為重要的網路服務之一。與部落格相關的名詞相當多，例如 Blogger、Blogroll 等。部落客(Blogger)通常被視為撰寫部落格文章的人或是部落格的所有者；而 Blogroll 則是由部落格的好友列表所串接起來的好友關係，將好友的部落格以網站連結方式加入至個人部落格中。

這種由部落格好友所衍生出來的關係稱之為線上社會網路(Online social network)。Newman 於 2003 年的研究中將社會網路(Social network)定義為「一群人或一個群體中的成員，透過某些模式的互動或聯繫所產生出的關聯就為社會網路。」[7]。目前與部落格線上社會網路有關的研究，根據其關係的型態可分作兩類：第一類的社會網路關係多以上述部落格之間的好友關係

作為主要的關係依據，這類關係稱之為**外顯關係**(Explicit Relationship)。除此之外，本研究認為可以透過文字分析與比較的方式，取得部落格之間隱含的關係，這第二類關係則稱為**內隱關係**(Implicit Relationship)。舉例來說，內隱關係的獲得可藉由部落格帳號、部落格站名、文章內文、文章分類等相關資訊來比較其相似度程度，或是透過部落客暱稱和他人對於部落客之稱謂來找到部落客之間的內隱關係。

Recuero 在 2008 年發表的研究論文中發現，該篇研究的研究對象平均擁有 2 至 2.5 個部落格 [8]，與目前認知部落格為單一部落客擁有的觀念有所不同。本研究的目的是之一，則是透過內隱關係的分析比較，找尋並證明單一部落客是否擁有 2 個或以上的部落格。

經找尋到單一部落客擁有多個部落格後，再透過部落格間的好友外顯關係，以及透過部落客稱謂所找出的內隱關係，合併兩者將部落客社會網路圖資訊形象化，並透過圖形分析方式，試圖探討出部落客在合併外顯與內隱後，是否會產生部落客間的關係變化，抑或部落客在整體社會網路中的變化以及是否會改變整體社會網路的型態。本研究於第二節將對部落格、社會網路與圖形理論進行文獻探討；第三節將說明本研究之實驗對象、方法與流程；第四節將說明本研究實驗結果與發現；第五節則對本研究進行結論。

二、文獻探討

由於近年來網際網路迅速發展，以及 Web2.0 概念 [10] 的提出，使得使用者大量參與網路相關服務，進而形成許多線上社群網站，亦形成一股線上社群交友的風潮。因此，與線上社群與社會網路的相關研究越來越多。

線上社會網路

目前有關線上社會網路的相關研究可分成

幾類：

- **論文資料庫作者社會網路**：這類研究藉由共同作者的姓名來探討出作者間的社會網路關係 [11]。
- **封閉式交友網站社會網路**：這類研究主要探討封閉式交友網站中人與人之間的結構拓撲分析 [3]。
- **部落格社會網路**：這類研究透過好友連結來探討出部落格之間的社會網路關係或透過文章內容來找出部落格間的社群關係 [1] [6]。

部落格的分類

資策會於 2006 年的研究中將部落格分成四類，其分類象限圖如圖 1 所示 [1]。

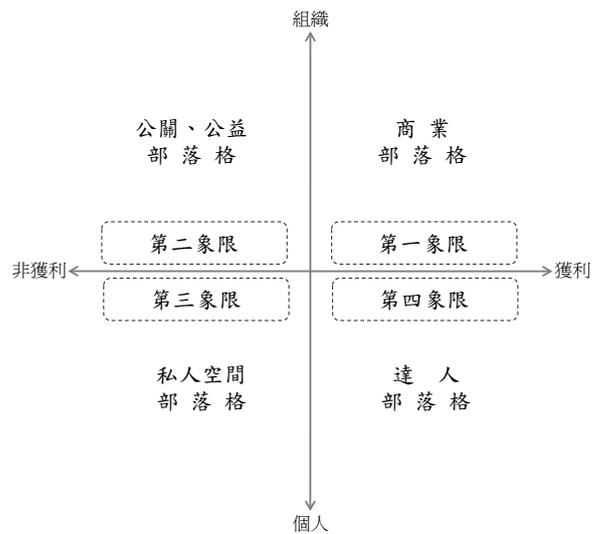


圖 1 部落格分類象限圖

- **商業部落格**：該類部落格以獲利當作經營目的，而再透過該部落格發展其他商業機制。
- **公益或公關部落格**：由於該類部落格多為非營利機構經營，且以溝通或公關、宣傳當作經營目的與重點。
- **私人空間部落格**：為個人部落格，常以個人抒發情緒、記錄日常生活與分享資訊為使用

方式。

- **達人部落格**：該類部落格之所有人，或許已經於現實社會或虛擬網路上已有不小名聲，經營這類部落格則為擴散個人品牌或名氣為目的。

圖形理論

圖形理論(Graph Theory)於電腦科學領域中，經常使用於網路的軟體配置與相關應用上。圖形(Graph)基本上是由點(Vertex)與邊(Edge)所組成，其定義為圖形 $G=(V,E)$ 是由一群 V 與 E 所組成， V 不能為 0 且 E 為一對 V (兩個 V)之間的連接線。 $V=\{v_1,v_2,v_3,\dots,v_N\}$ 可視為一群節點或個體； $E=\{e_1,e_2,e_3,\dots,e_K\}$ 可視為一群關聯、關係 [7]。

三、研究方法

本研究為了探討單一部落客擁有多少部落格，與部落客社會網路於合併內隱關係之後是否有差異。因此，本研究實驗設計流程如圖 2 所示，其流程分成實驗對象取得、相似度分析、合併部落客、找尋部落客內隱關係與社會網路形象化等五階段進行。

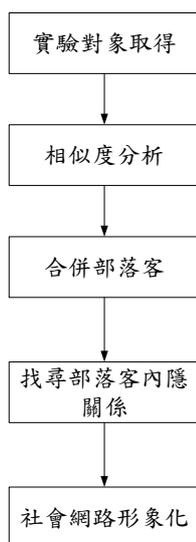


圖 2 實驗流程圖

實驗對象取得

本研究的實驗對象取得共分成兩階段，第一階段於 2009 年 6 月 16 日取得無名與痞客邦各五個分類的第一名部落格帳號(如表 1 所示)，並取得其部落格的好友列表與部落格相關資訊。

表 1 無名與痞客邦各五類之部落格列表

BSP	帳號	BSP	帳號
痞客邦	alicebrownie	無名	bajenny
痞客邦	abig99	無名	off60
痞客邦	terisawu	無名	illyqueen
痞客邦	anyichen	無名	sana217
痞客邦	cwyuni	無名	terisawu

為了使社會網路圖不皆為單向關係，故進行第二階段取得，本階段針對第一階段的好友進行其好友的取得，以利其社會網路關係為雙向關係，以使其更能顯示整體社會網路。

相似度分析

本研究針對實驗對象的帳號、部落格站名、文章分類與文章等四類，進行相似度比對，期比對方式如下：

- 帳號、相似度比對則採用字串比較方法，其用以比對字串是否完全相似，若相似則為 1，不相似則為 0。
- 文章分類本研究採用字串比較方法，並以 Jaccard's Coefficient 的方式來計算整體的文章分類相似度。Jaccard's Coefficient 常被用以測量兩向量間的相似程度，當數值越接近 1 時，則表示兩者間的相似程度越高，反之則越不相似 [4]。
- 部落格站名相似度比對本研究採用 Jaccard's Coefficient。
- 文章相似度比對則採用餘弦相似度 (Cosine Similarity)，該相似度比對最常使用於衡量文件之間的距離衡量方法，其以兩文件的關鍵字向量計算，並計算出兩文件間的餘弦夾角數值，其數值介於 0~1 之間。當夾角數值越接近 1 時，則代表兩文件相似度越高；反

之，則相似度越低。本研究因考慮到發布日期之因素，故於比對兩篇文章之相似度時，若其發布日期不同，則不予分析其相似程度。

合併部落客

本階段根據上一步驟的相似度分析結果，找到部落格與部落客間的關係。依據帳號、部落格站名、文章分類與文章相似結果，找到哪些部落格可以合併成單一部落客。本研究的文章相似度篩選條件是取得以該篇文章於比較對象的全部文章中相似程度最高之文章。再將合併後的部落客所持有的部落格好友轉成部落客的好友，已轉化成個人的好友列表。

找尋部落客內隱關係

於合併相似部落格於同一部落客後，本研究使用先前所建立的無名好友系統，利用帳號方式取得其他部落客對該部落客的稱謂。但為了使實驗的稱謂更加準確，故本研究再將稱謂輸入至無名挖挖挖(<http://www.mining.tw>)網站作修正。無名挖挖挖網站是以關鍵字搜尋個人好友的無名部落格，其結果會顯示擁有相同稱謂的部落格，供使用者自行選取。故本研究則利用該網站的關鍵字搜尋而得到準確的稱謂，最後再以關鍵字搜尋部落格文章，取得其部落客於文章中的內隱關係。

社會網路形象化

UCINET6 為目前社會網路分析的主要分析工具之一，該軟體不僅可將社會網路關係資料轉化成社會網路關係圖，亦提供計算節點間的距離與重要程度[9]。本階段將根據上列實驗步驟所得之社會關係資訊，使用 UCINET6 社會網路分

析工具繪製成外顯、內隱與合併外顯與內隱的社會網路關係圖，並再依此社會網路關係圖進行圖形分析。

四、研究結果

本研究實驗結果從 25 個部落格中，透過帳號與站名相似度分析找出 8 個部落格極為相似，並將這 8 個部落格分成 4 組進行文章分類與文章的相似度分析。從結果發現有 8 個部落格可合併成 4 位部落客，其結果如表 2 所示。這 8 個部落格的站名與帳號相似度皆達 1；其文章分類相似度最低達到 85% 相同，最高則達到 100% 完全相同；文章相似度部分，其相似度最低僅達 21%，最高也達 63% 相似度，故根據上列相似度比對結果，將其合併成 4 位部落客。

本研究經實驗後，得到利用好友連結所產生的部落客外顯社會網路關係圖（如圖 3）、透過部落客稱謂來搜尋文章內文所產生的部落客內隱社會網路關係圖（如圖 6）以及合併內隱與外顯的社會網路關係圖（如圖 7）。

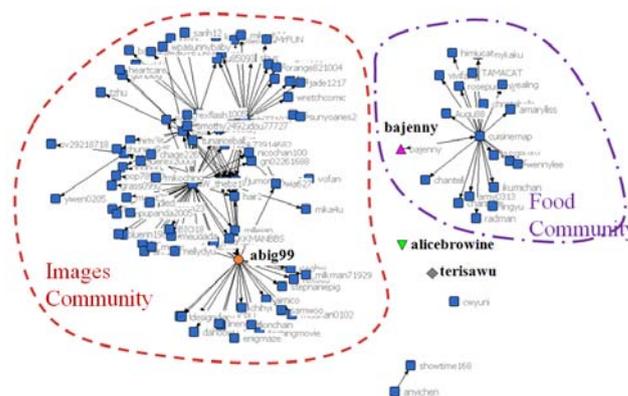


圖 3 部落客外顯社會網路圖

表 2 部落客合併結果列表

痞客邦部落格帳號	無名部落格帳號	帳號相似度	站名相似度	文章分類相似度	文章相似度	好友關係	合併後的部落客名稱
alicebrownie	alicebrownie	1	1	99%	46%	好友關係	alicebrownie
bajenny	bajenny	1	1	100%	21%	好友關係	bajenny
abig99	abig99	1	1	86%	42%	無好友關係	abig99
terisawu	terisawu	1	1	85%	63%	無好友關係	terisawu

圖 3 為以好友列表所組成的外顯部落客社會網路圖。本研究合併部落格後所得的部落客，以不同於正方形的形狀標示出來。由圖 3 中可發現其社群主要分成兩大群體，依據其部落格的文章內文可得知左邊的社群主要由圖像創作為主的部落客所組成；右邊群體以美食介紹、討論為主的部落客。

圖 4 為整體部落客內隱社會網路圖。把內隱社會網路與外顯社會網路結合之後做比較，可發現許多有趣的觀察點。以 cwwany 為例，若單以外顯討論其社會網路關係，其結果就如同圖 5 所示，同中圓形圖示為 cwwany。透過發掘其內隱關係後，可發現 cwwany 認識圖 5 的其他部落格，而且其內隱關係比外顯關係更為複雜且豐富，其 cwwany 的內隱關係如圖 6 所示，其相互連結的部落客以三角形圖形表示。

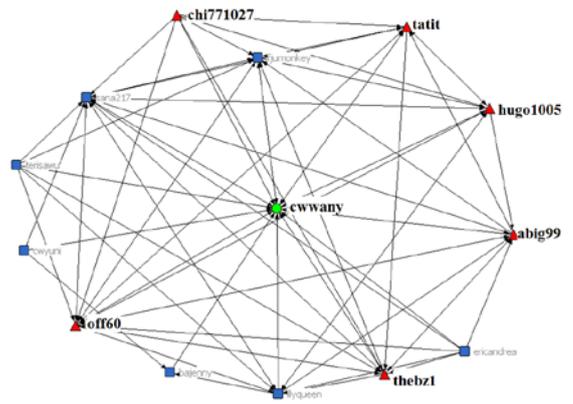


圖 6 cwwany 內隱社會網路圖

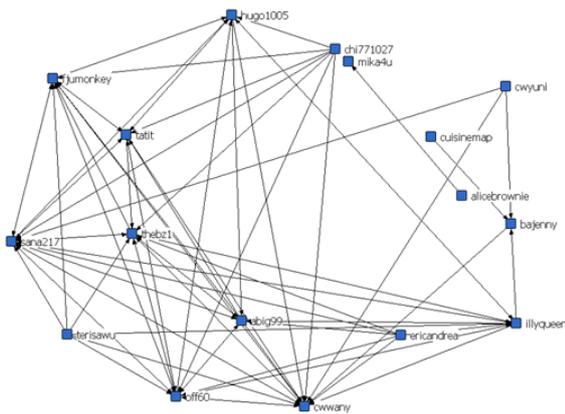


圖 4 部落客內隱社會網路圖

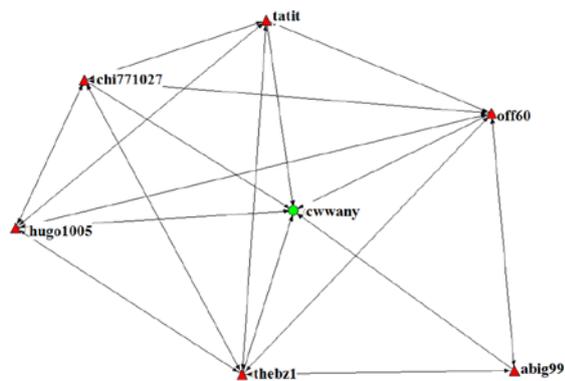


圖 5 cwwany 外顯社會網路圖

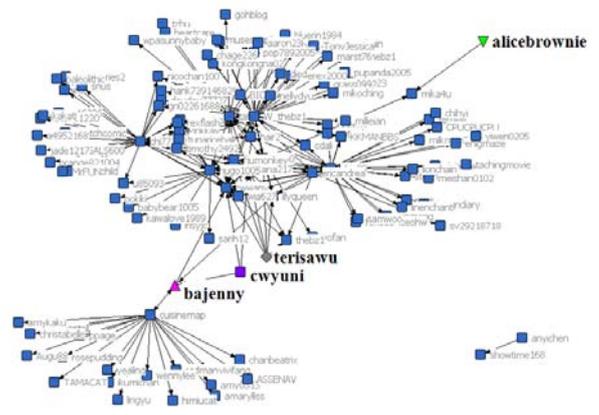


圖 7 合併後之部落客社會網路圖

在外顯部落客社會網路圖(如圖 3)，加入內隱社會網路圖(如圖 4)，可獲得如圖 7 的合併社會網路圖。相較於圖 3 的外顯部落客社會網路圖，可發現原本圖 3 裡的圖像創作與美食兩個群體，經加入內隱關係，在圖 7 中兩群體間產生關聯。兩群體因加入 bajenny 的內隱關係，進而將兩個群體合併成單一的社會網路。因 bajenny 為兩群體中的重要橋樑，亦提升 bajenny 於整個社會網路中的重要程度。

此外，圖 3 中原無與其他部落客產生關係的獨立部落客 alicebrownie、terisawu、cwyuni，在加入內隱關係後與群體產生關聯。

五、結論

由於部落格擁有個人化、服務整合等特性，進而受到社會大眾、公司團體等的喜愛，以致使

部落格發展出線上社會網路關係。本研究提出除了傳統的好友外顯關係之外，線上社會網路可透過自然語言處理的文章比對等技術，分析比較獲得部落格之間的內隱關係。本研究透過內隱關係，用來偵測部落格是否為單一部落客所擁有。此外，本研究亦試著探討出部落客間的內隱關係，並以形象化方式形成各社會網路圖，並依此圖形進行探討與討論。經本研究實驗後發現下列幾點：

- 根據本研究結果，發現到有四位部落客各擁有兩個部落格，與 Recuero 的研究中敘述結果相符。
- 外顯與內隱部落客社會網路關係確實不同，內隱關係可能隱含更多、更深入且更多元的社會關係。
- 合併外顯與內隱的社會關係後，會造成部分獨立的部落客的關係產生連結。
- 外顯關係在加入內隱關係後，使得原本兩個不同的群體產生關聯，並且使部分部落客的社會關係的重要程度提升。

本研究目前僅以帳號加稱謂方式找尋部落客間的關係，未來本研究將繼續找出更多不同面向的內隱與外顯社會關係，並再進一步於語意分析、智慧型機器人方向邁進，以便更深入的探討出部落格、部落客間的其他相似、關聯之處。

六、參考文獻

- [1] 陳信宏，周樹林，“Blog 獲利方式與案例分析”，資策會 MIC，民國 95 年。
- [2] C. Marlow, “Audience, structure and authority in the weblog community.”, Paper presented at the International Communication Association Conference., 2004.
- [3] F. Fu, L. Liu, L. Wang, “Empirical analysis of online social networks in the age of Web 2.0”, *Physica A: Statistical Mechanics and its Applications*, Vol. 387, pp. 675-684, 2008.
- [4] G. Salton, A. Singhal, M. Mitra and C. Buckley, “Automatic text structuring and summarization”, *Information Processing & Management*, Vol. 33, No. 2, pp.193-207, 1997.
- [5] J. Barger, “Weblog Resources FAQ”. www.robotwisdom.com/weblogs, 1997.
- [6] M. Chau and J. Xu, “Mining communities and their relationships in blogs : A study of online hate groups”, *Int. J. Human-Computer Studies*, Vol. 65, Issue 1, pp.57-70, 2007.
- [7] M. E. J. Newman, “The structure and function of complex networks.”, *SIAM Review*, Vol. 45, No. 2, pp.167-256, 2003.
- [8] R. C. Recuero, “Information Flows and Social Capital in Weblogs: A Case Study in the Brazilian Blogosphere”, *Conference on Hypertext and Hypermedia Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pp.97-106, 2008.
- [9] S.P. Borgatti, M.G. Everett and L.C. Freeman, “Ucinet for Windows: Software for Social Network Analysis”, Harvard, 2002.
- [10] Tim O'Reilly, “web1.0 to web2.0”, <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- [11] X. Liu, J. Bollen, M. L. Nelson, H. V. d. Sompel, “Co-authorship networks in the digital library research community”, *Information Processing & Management*, Volume 41, Issue 6, Special Issue on Infometrics, pp.1462-1480, 2005.