

國立交通大學

多媒體工程研究所

碩士論文

複雜網路中橋接式與強鍵式連結之偵測、

分析與應用

Bridge and Bond Link in Complex Networks:

Detection, Analysis and Applications

研究生：黃全榮

指導老師：孫春在 教授

中華民國 九十九 年六月

複雜網路中橋接式與強鍵式連結之偵測、分析與應用
Bridge and Bond Link in Complex Networks: Detection, Analysis and
Applications

研究生：黃全榮

Student : Chuan-Jung Huang

指導老師：孫春在

Advisor : Dr. Chuen-Tsai Sun

國立交通大學
多媒體工程研究所
碩士論文

The logo of National Chiao Tung University is a circular emblem with a gear-like border. Inside the circle, there is a stylized building and the year '1896'. The text 'NCTU' is also visible within the emblem.

A Thesis
Submitted to Institute of Multimedia Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
In
Computer Science
June 2010
Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

複雜網路中橋接式與強鍵式連結的偵測、分析與應用

學生：黃全榮

指導教授：孫春在 博士

國立交通大學

多媒體工程研究所

中文摘要

強弱連結是分析複雜網路的核心概念之一。然而，先前的研究對於網路連結強弱的定義面臨兩種挑戰：一、對於模糊不清的強弱定義，在應用上認定困難。二、雖然嚴格定義了強弱的界線，卻以非強即弱的判斷方式區分連結性質，未能反映許多真實網路的性質。我們強調複雜網路具有階層性質，並認為複雜網路中的連結在不同階層時具有不同的特性，有些連結對於區域訊息傳遞具有較大的影響力，但是對於全域訊息傳遞的影響力較低；有些連結對於區域到全域的訊息傳遞都極具影響力。

為了避免在量化連結性質時模糊不清的強弱定義或非強即弱的判斷方式，我們基於連結特性提出了橋接式與強鍵式連結的概念，並且設計一套偵測演算法，該演算法能夠依據拓撲與統計上的意義偵測出某特定連結在不同網路階層中所具有的性質。在分析之後，我們將連結分為下列四種類型：全域橋接式連結、N階橋接式連結、強鍵式連結和沉沒橋接式連結。以此為基礎，我們觀察三種不同類型真實網路中連結在不同階層時具有的性質，瞭解網路中每一條連結從區域觀點到全域觀點的性質變化，並且依據演算法將連結分成不同類型。藉由這些不同

類型的連結在網路中存在的比例，我們能夠描述網路的拓樸性質與動態特性。

最後，我們以連結的階層性質與連結類型分佈比例為基礎提出了三個應用：
一、尋找網路中的社群結構。二、真實網路分群。三、分析真實網路與理論網路在拓樸架構上的相似性。實驗結果顯示，本研究提出的概念除了在這些應用上得到不錯的結果之外，更能夠以一致的概念與方法解決不同的問題，說明了掌握連結的階層性質能夠給予研究學者在網路拓樸架構的建構或演化原理上更深入的瞭解。

關鍵詞：強弱連結、複雜網路、分群、社群結構



Bridge and Bond Link in Complex Networks: Detection, Analysis and Applications

Student : Chuan-Jung Huang Advisor : Dr. Chuen-Tsai Sun

Institute of Multimedia Engineering
National Chiao-Tung University

Abstract

A concept of strong-ties and weak-ties played an important role in complex network studies. However, different researcher would face following problem in defining strong or weak of links: (1) Ambiguous definition between strong and weak, and (2) Clear definition of limit between strong and weak but link is either strong-tie or weak tie (dichotomy method). We considered that characteristic of links varies in different network layers (hierarchy property). Some links have obvious influence on information transmission in local area, but have no effect on information transmission in global scale; and some link have significant influence on information transmission from local to global scale.

To avoid ambiguous definition and dichotomy method when quantifying links, we applied characteristic of links to propose the concept of bridge link and bond link, and designed an algorithm which can detect varies characteristic of links in the different network layers according to topology properties and statistically significant, then classify links into four types :

Global bridge link, N degree bridge link, Bond link, and Sink bridge link. In this paper, we observed the characteristic of links in different layers under three real networks, and realized variation in properties of every links from local to global view, then classified links into different types according to the algorithm. We can describe the topology structure and specific dynamic properties of networks by means of the proportion of different link types.

Ultimately, we proposed three applications based on hierarchy property of links and the proportion of different link types: (1) find community structure in complex network; (2) complex network clustering; (3) analyzing the correlation between real-world network and theoretic network on topology structure. The results demonstrated that the concept we proposed not only led to decent results on these applications, but also resolved different problems with the consistent methodology. By analyzing and realizing hierarchy property of links, researchers in many fields of science will have a better understanding of design principle or evolution principle beyond complex network.

Keywords : strong-ties and weak-ties, complex network, clustering, community structure.

誌 謝

在許多人的協助之下完成了這篇研究論文，首先要感謝我的母親總是默默的支持我與鼓勵我，讓我在兩年的碩士生涯中沒有後顧之憂。接著要感謝我的指導教授孫春在先生，在孫老師的帶領下，我學習到什麼是「作研究」，使我瞭解一篇論文不只是在解決研究問題，更是要瞭解研究的動機和研究的重要性與貢獻，並且能夠完整的呈現一篇論文的價值所在。然後要感謝我的老師兼學長黃崇源先生，從研究問題的產生、實驗方法與數據分析都給了我非常多的建議與指導，除了研究上的指導之外，在人生規劃與生活態度上也給予我許多方向與意見，讓我更清楚未來的目標。

感謝口試委員項潔教授、曾憲雄教授以及陳穎平教授，在口試的過程中給予我許多建議與指導，使得本篇研究論文能夠趨於完善。

感謝實驗室的學長，宇軒、聖文、王豪、宜睿、基成和立先，在口試投影片的呈現與製作上給予了很多意見，使得我在口試報告時能夠更事半功倍。也感謝實驗室碩二的同伴們，壯為、嘉宏、璽文、謹譽、泰源和柏志，大家在作研究的日子裡互相鼓勵與協助，是一群不可缺少的好夥伴。也要感謝實驗室的學弟妹們，偉存、誌宏、承宏、景照、皓琮和順貞，無論是在課業上或是生活上的幫忙，能讓我省下了許多時間，可以將心力完全放在研究上。感謝學習科技實驗室的每一位同學，讓實驗室的氣氛非常的融洽，是完成論文寫作的原動力。在此，致上最深的謝意。

目錄

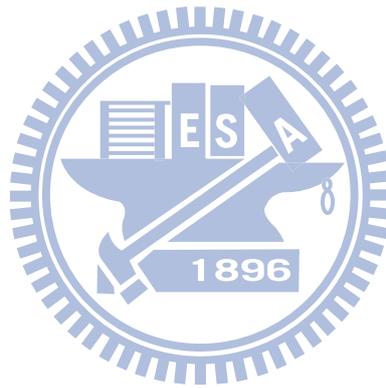
中文摘要.....	i
Abstract.....	iii
誌謝.....	v
目錄.....	vi
表目錄.....	viii
圖目錄.....	ix
第一章 緒論.....	1
1.1 研究背景.....	1
1.2 研究動機.....	4
1.3 研究目標.....	6
1.4 研究重要性.....	7
第二章 文獻探討.....	9
2.1 拓撲特性影響複雜網路的行為.....	9
2.2 強弱連結.....	15
2.3 從局部觀點到整體觀點.....	20
2.4 Edge-betweenness 與連結兩端節點的共同朋友.....	22
第三章 研究方法.....	24
3.1 實驗目的與概念描述.....	24
3.1.1 連結的「階層」(Hierarchy).....	26
3.1.2 連結兩端節點的共同朋友.....	28
3.2 實驗設計.....	30
3.2.1 產生「對應的隨機網路」.....	30
3.2.2 判斷連結種類.....	31
3.2.3 討論.....	36
第四章 研究結果與發現.....	37
4.1 理論網路.....	37
4.1.1 WS-小世界網路.....	37
4.2 真實世界網路.....	42
4.2.1 社會網路(朋友網路).....	42
4.2.2 生態網路.....	44

4.2.3 電子電路網路.....	47
第五章 演算法穩固分析與應用.....	50
5.1 穩固分析.....	51
5.2 演算法應用.....	56
5.2.1 複雜網路的分群.....	56
5.2.2 找尋網路中的社群結構(Community structure)	61
第六章 結論.....	67
第七章 未來展望.....	70
附錄.A	72
附錄.B.....	73
附錄.C.....	74
附錄.D	75
附錄.E.....	76
附錄.F.資料來源	83
參考文獻.....	84



表目錄

表 1：四種社會網路的拓樸資訊.....	43
表 2：七種食物鏈網路的拓樸資訊.....	45
表 3：三種電子電路網路的拓樸資訊.....	48
表 4：不同連結類型性質整理表.....	68
表 5：不同繞線機率下的 WS 小世界網路	76
表 6：不同繞線機率下的 NW 小世界網路	82



圖目錄

圖 1：連結的影響範圍。.....	4
圖 2：連結種類。.....	6
圖 3：13 種大小為三節點的基調。.....	12
圖 4：產生對應的隨機網路.....	12
圖 5：社會網路中的強弱連結。.....	15
圖 6：連結強弱與資料複雜度的關係圖。.....	17
圖 7：整體中心度示意圖。.....	21
圖 8：節點在網路中的位置與中心度關係圖。.....	22
圖 9：共同朋友、edge-betweenness 與實際通話量之間的關係。.....	23
圖 10：階層概念圖。.....	27
圖 11：節點 i 延伸一步之後所拜訪到的節點形成 i_1 。.....	28
圖 12：共同朋友。.....	28
圖 13：共同朋友概念圖。.....	29
圖 14：實驗流程圖。.....	30
圖 15：交換演算法示意圖。.....	31
圖 16：演算法流程圖.....	33
圖 17：WS-Small world, $N=100, \langle k \rangle=4$ 。.....	38
圖 18：WS-Small world, $N=100, \langle k \rangle=8$ 。.....	38
圖 19：WS-Small world, $N=100, \langle k \rangle=16$ 。.....	39
圖 20：WS-Small world, $N=100, \langle k \rangle=32$ 。.....	39
圖 21：WS-Small world, $N=100, \langle k \rangle=64$ 。.....	39
圖 22：四種社會網路連結種類分佈結果統計圖。.....	43
圖 23：七種食物鏈網路連結種類分佈結果統計圖。.....	46
圖 24：Trophic food chain(TFC)。.....	46
圖 25：Bridge Brook Lake 的食物鏈網路示意圖。.....	47
圖 26：三種電子電路網路中連結種類分佈結果統計圖。.....	48
圖 27：s208 的電子電路網路拓撲架構示意圖。.....	49
圖 28：演算法判斷失真網路的比對正確率統計圖。.....	51
圖 29：移除連結事件下，「一階強鍵式連結」的比對正確率。.....	54
圖 30：移除連結事件下，「全域橋接式連結」的比對正確率.....	55
圖 31：移除連結事件下，「其他連結」的比對正確率.....	55
圖 32：四種社會網路在分群指標下的相關性($C=0.9\sim 0.99$)。.....	58
圖 33：七種食物鏈網路在分群指標下的相關性($C=0.62\sim 0.99$)。.....	58
圖 34：三種電子電路網路在分群指標下的相關性($C=0.7\sim 0.92$)。.....	58
圖 35：14 種網路彼此之間的相關係數矩陣視覺圖。.....	59

圖 36：14 種複雜網路與理論網路之間的相關係數矩陣視覺圖。	60
圖 37：具有社群結構的簡單網路。	61
圖 38：Zachary 研究美國某大學空手道俱樂部成員互動網路。	63
圖 39：GN 演算法分析 Zachary 網路的結果。	63
圖 40：經由找尋網路中社群架構步驟所得到的樹狀圖。	64
圖 41：本研究所提出的演算法分析 Zachary 網路的結果。	65
圖 42：共同朋友與連結方向性。	70
圖 43：共同朋友與連結權重值。	71
圖 44：傳播速率圖。	73
圖 45：不同性質的社會資本混合下所產生不同型態的社會。	74
圖 46：共同朋友比例。	75
圖 47：WS-small world, $N=100, p=2\%$ 。	77
圖 48：WS-small world, $N=100, p=4\%$ 。	77
圖 49：WS-small world, $N=100, p=8\%$ 。	78
圖 50：WS-small world, $N=100, p=16\%$ 。	78
圖 51：WS-small world, $N=100, p=32\%$ 。	78
圖 52：WS-small world, $N=100, p=64\%$ 。	79
圖 53：NW-small world, $N=100, p=2\%$ 。	80
圖 54：NW-small world, $N=100, p=4\%$ 。	80
圖 55：NW-small world, $N=100, p=8\%$ 。	80
圖 56：NW-small world, $N=100, p=16\%$ 。	81
圖 57：NW-small world, $N=100, p=32\%$ 。	81
圖 58：NW-small world, $N=100, p=64\%$ 。	81

第一章 緒論

1.1 研究背景

複雜性科學是當前科學發展的焦點與前沿，其理論研究與實務應用正在向各個學科領域滲透，成為一門蓬勃發展且受到眾多科學家關注的跨領域學科。複雜系統涉及的範圍非常廣闊，包括自然、工程、生物、經濟、管理、政治與社會等各個方面；從一個細胞呈現出來的生命現象，到經濟股票市場的漲落、城市交通運輸的管理、自然災害的預測與預防，乃至社會國家的興衰等等。由於各個學科領域對於複雜性的認識和理解都不大相同，所以學術研究人員為了避開專有名詞的爭論，遂採用「複雜系統」來代表在變化無常的活動中，所浮現出的某種秩序，其中的演化、湧現、自組織、自適應、自相似現象被認為是複雜系統共同且重要的特徵。

複雜系統是一個由多個簡單單元所組成的結構。一般來說，複雜系統最特別的地方就是這些簡單單元之間會經由非線性的交互作用，而產生集體的行為。為了瞭解複雜網路中個體互相影響之後產生的集體行為，許多統計上和數學上的工具已經被發展出來，進而幫助複雜系統的研究。在過去幾年的研究當中對於複雜系統的瞭解有重大的發展，其原因有：第一、資料收集或取得容易，可建立大型資料庫；第二、電腦的計算能力日益強大，能夠將收集的資料做大型的模擬；第三、跨領域的研究使得各個領域互通方便，能夠互相存取研究結果；第四、越來越多簡化的方法，能夠去瞭解整個複雜系統的動態行為。

複雜系統中的集體行為可以在空間或時間上表現為圖形化的結構。網路是現代複雜科學中重要的研究工具之一，許多複雜系統能夠簡化成網路的方式展現，使我們能夠藉由複雜系統的拓撲形式去瞭解個體間互相影響後所產生的集體行為。譬如說，當在研究社會現象或是瞭解輿論和八卦是如何在社會上傳播時，我們可以將人簡化成節點，人與人之間的社交關係則簡化成為連結；當在研究某個

地區的生態系為何發生物種間的消長或滅絕時，我們可將物種簡化成節點，物種間食物鏈的關係簡化成連結；當在研究網際網路架構如何影響資訊傳播路徑或是網路安全時，我們可將路由器簡化成節點，實體線路則簡化成連結。不同的複雜系統都有各自不同的意義，但是當其表示成複雜網路時，都能夠以拓樸的角度來探討其所發生的各種現象，因此複雜網路本身已經成為新興科學中相當重要的研究領域。

複雜網路中最重要基礎元素就是連結。連結代表了複雜網路中個體間資訊流動的途徑，而這些途徑勾勒出複雜網路所呈現的拓樸形式會直接影響到複雜網路的動態行為，舉例來說，複雜網路的分隔度、群聚度與階層結構都會影響到複雜網路所產生的現象 (Clauset, Moore & Newman, 2008; Dakker, 2007; Newman, 2003)，網路中存在具有功能性的基調(motifs)對於網路中的訊息傳遞提供了合適的拓樸架構(Artzy-Randrup, Fleishman, Ben-Tal & Stone, 2004; Huang, Cheng & Sun, 2007; Milo et al., 2002; Milo et al., 2004)。Granovetter 提出了弱連結(weak tie)與強連結(strong tie)的概念，他的『找工作』實驗說明了弱連結對於複雜網路中資訊的傳遞扮演了極為重要的角色(Granovetter, 1973; Papakyrizis & Boudourides, 2001)，它能夠對於網路中的群體帶來新的資訊以及將資訊傳到更廣的範圍；此外，某些研究也指出弱連結對於複雜網路有穩定性的作用(Csermely, 2006; McCann, Hastings & Huxel, 1998)，當網路發生危難或是被攻擊，只要大部分弱連結還能夠存在，整體網路還是能夠維持運作而不至於崩潰。相反地，強連結對於資訊的傳遞較為不重要，但是對於複雜網路中的群體或社群中個體間的資訊交流有著重要的影響。這些強連結能夠維持群體中個體間的緊密關係，個體不會輕易的消失或離開於此群體。舉例來說，社會網路上充滿了許多群體，學校中的班級、公司中的部門或是國家內的政黨等等，這些群體內的成員彼此間都有共同的目標、話題甚至是文化。強連結的作用在於將這些共同的資訊保留在群體中，維持群體中的同質性。

類似強弱連結的概念也在 Putnam 提出橋接型社會資本和強鍵型社會資本中找到類似的看法(Putnam, 2000)，他認為橋接型社會資本屬於外展型，能夠包含分屬不同全體的人們，能夠為群體帶來新的資訊與機會。另外，強鍵型社會資本則屬於內聚型，傾向於強化排他性認同與同質性團體，群體中的成員彼此間能夠提供較多支持與鼓勵。對於家族企業的公司或是集團中，若是能夠擁有強鍵型社會資本，則個體更有機會在團體中獲得較好的職位或是職薪。

Hansen 利用了社會網路中強弱連結的概念與知識複雜程度的觀念去探討組織中的各個子單位要交換或分享知識時，單位間的連結特性會造成什麼樣的影響(Hansen, 1999)。他的研究發現組織中無論是強連結或是弱連結都能夠導致這些知識能夠有效的分享於單位之間，強連結和弱連結在知識分享時，對於知識的搜尋與傳遞均有優缺點。當知識的複雜程度較高時，弱連結會造成知識在傳遞時的阻礙，但是無論知識的複雜程度高或低，弱連結均會降低知識的搜尋時間。反觀強連結對於知識的搜尋沒有直接的幫助，但是當知識的複雜程度較高時，能夠有效的傳達知識到目的地。因此，經由這些研究結果我們可以知道強連結和弱連結在複雜網路當中各自擁有存在的功能與目的，對於網路來說兩種類型的連結都有其重要性。

以往在研究複雜網路時，大部分主要是聚焦在節點的特性上，想要瞭解哪一些節點居於網路核心的位置並且對於整體網路最具有影響力。值得注意的是，我們能夠從節點的區域特性(節點的分支度,degree)向外分析到全域特性(節點的接近性,closeness)。但是對於網路中節點間的溝通管道—連結，卻沒有從區域到全域式的探討。因此，在這篇研究論文當中將提出一套演算法能夠以不同網路尺度的觀點來分析網路中的連結，藉此瞭解每一條連結在網路中所扮演的角色。

1.2 研究動機

在以往研究中鮮少對於複雜網路中的強弱連結有清楚地定義，更重要的是對於強弱連結的區分也總是非強即弱的二分法。但是由網路層級的觀點來看，連結在不同層級的網路觀點中所具有的強弱性質也有所不同。也就是說，有些連結對於複雜網路中某個區域的訊息傳遞相當重要，但是對於整體複雜網路的資訊傳遞的影響卻不大；有些連結從區域到全域的訊息傳遞都具有相當大的影響力。舉例來說，以圖 1 中的網路為例，對於圖 1.(a)來說，此網路的藍色連結對於訊息傳遞有較大的影響力。但當考慮圖 1.(b)的多層級網路時，藍色連結的影響力限制在綠色或黃色節點所形成的子網路當中，而紅色連結卻對於整體網路的資訊傳遞擁有決定性的影響，一旦紅色連結消失，那麼黃色節點的資訊永遠不能傳遞到綠色節點，使得兩個子網路完全分離。另外，(a)、(b)兩圖中的黑色連結對於訊息傳遞都不具太大的影響力。

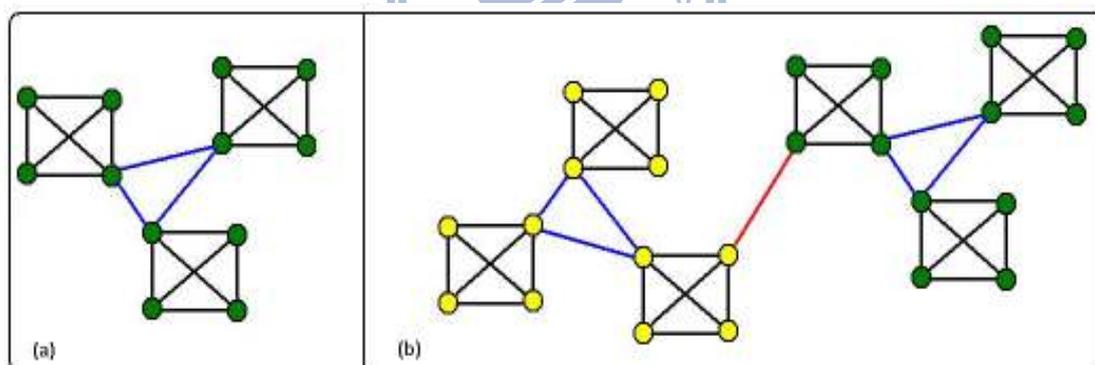


圖 1：連結的影響範圍。

因此複雜網路中的連結所扮演的角色會隨著網路層級的不同而有所不同，並不是以往認為的非強即弱的這種二分法。此外，許多研究已經證明了複雜網路是具有階層(hierarchy)性質的(Clauset et al., 2008)，然而，以某連結為中心向外延伸之後，此連結在不同階層之下的影響力是會消退或是增長，這樣的問題仍需要被解決(Goyal, 2008)。

過去對於複雜網路的研究當中，若是節點之間有連結，都會參考此連結背景

特性(domain-specific)的意義，並且找尋不同的方式量化這些特性(Barrat, Barthélemy, Pastor-Satorras & Vespignani, 2004; Papakyriazis & Boudourides, 2001; Onnela et al., 2007)。舉例來說，雖然性關係網路和生態網都是複雜網路，但是在性關係網路中，兩節點若有性關係，那麼就擁有連結；但是想要量化此連結的背景特性，可能就要紀錄兩節點在一定時間內的性交次數。對於生態網路來說，兩節點之間若是有“捕食者—獵物”的關係，那麼就擁有連結；但是想要量化此連結背景特性，可能就要紀錄捕食者在一定時間之內捕食獵物的數量。對於這兩種複雜網路來說，除了量化方式大不相同之外，這些紀錄資料的取得或蒐集都很困難，而且最重要的是這些資料的正確性也難以驗證。

藉由複雜網路的拓樸特性，我們能夠由複雜網路最底層的觀點探討連結的強弱。更重要的是，複雜網路的拓樸特性使我們能夠使用相同的方式找出不同複雜網路間共同的特性，目前複雜網路所發現的共同全域拓樸特性有三點：網路平均路徑短(Small world property)、群聚性高(Highly clustered connection)以及節點 degree 呈現冪次率分佈(Scale-free phenomenon)(Albert & Barabási, 2002; Goh, Oh, Jeong, Kahng & Kim, 2002; Watts & Strogatz, 1998; Strogatz, 2001)。經由這些拓樸特性我們可以不考慮複雜網路的背景物理意義，而直接透過研究拓樸架構得到這些與網路行為息息相關的拓樸特性；另外更重要的是，這些共同特性說明了複雜網路的拓樸架構並不是“偶然”產生的，複雜網路的拓樸架構擁有某種設計原則或演化原則，經由上一節也說明拓樸特性對於複雜網路的行為有著重要的影響性，因此，我們希望利用拓樸的觀點來研究複雜網路中連結的特性。

1.3 研究目標

在本研究中，為了要避免強弱連結模糊不清的定義，我們提出橋接式連結(bridge link)以及強鍵式連結(bond link)的概念。橋接式連結代表網路中兩個群體間聯繫的力量，它對於遠距離的訊息傳遞有很大的影響力；強鍵式連結則代表了一個群體中兩個體間的聯繫力量，它對於個體間的互助資源有較大的影響力。因此，橋接式連結與強鍵式連結在複雜網路中扮演了不同的角色，其影響也大不相同。根據橋接式/強鍵式連結的概念我們提出一個演算法，它能夠根據拓樸特性判斷出連結在不同網路層級時所具有的性质，給予橋接式/強鍵式連結拓樸與統計上的定義。此外我們也依據連結的特性分為：沉沒橋接式連結(Sink bridge link)、N 階橋接式連結(N degree of bridge link)、全域橋接式連結(global bridge link)以及強鍵式連結(Bond link)，代表了此連結對於不同層級所扮演的角色，能夠讓我們瞭解連結從區域到全域的影響性。全域橋接式連結的存在對於整體網路的訊息流動或是能量移轉的範圍具有最大的影響力，如圖 2 中的紅色連結(\overline{DE})。強鍵式連結在網路中的群體內呈現較為緊密的出現，對於群體內同質化有加強的作用，如圖 2 中的黑色連結。N 階橋接式連結對於網路的功能介於全域橋接式連結與強鍵式連結之間，如圖 2 中的藍色連結(\overline{AC} , \overline{BC} , \overline{FG} , \overline{FH})。沉默橋接式連結是網路中邊緣節點和整體網路連接的邊，如圖 2 的橘色連結(\overline{BI} , \overline{JG})。

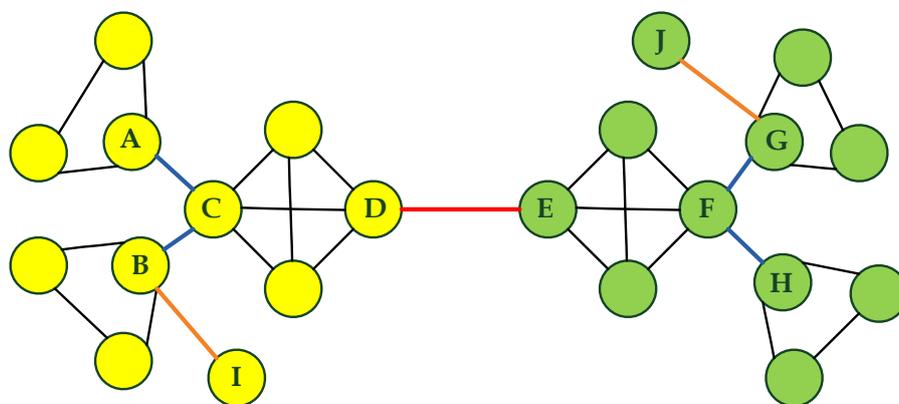


圖 2：連結種類。

本篇研究論文中也分別針對各種理論網路與實際網路，觀察橋接式/強鍵式連結對於群聚度或分隔度等拓樸特性的影響，能夠更進一步的瞭解各種類型的連結在這些網路中具有的功能與扮演的角色。

1.4 研究重要性

無論是何種領域的複雜網路，個體展現的行為或是群體發生的現象都會和其所嵌入的網路形式脫離不了關係。舉例來說，由社會學的觀點來看，我們自己所建立的人際關係網路會影響到我們生活的各方面，包括了健康狀況、職業選擇或是個人身份。研究指出，缺乏社會聯繫與社區聯繫的人與接觸面廣泛的人相比，死亡率更高(Berkman & Syme, 1979)；處於團體之間的溝通者或是人際關係良好的人，通常擁有較高的社會地位。由生態網路的觀點來看，擁有越多食物鏈關係的物種，在遭遇環境劇變時，能有較高的存活率；缺少食物鏈關係的物種，很可能因為其他少數物種的遷移或是消失而面臨絕種的危機。就政策制訂而言，利用複雜網路的拓樸分析能夠預測組織將要發生的現象並且提早做出防範與規劃。複雜網路中個體間的關係形式更是我們想要探討的重點，我們相信個體之間的相互作用會直接影響個體本身的行為，因此，本研究探討複雜網路中連結的特性，藉此看到不同性質的連結在網路中具有什麼樣的影響力。

一旦瞭解了連結從區域到全域對於複雜網路的影響，那麼針對不同的複雜網路，能夠依據其中連結重要性的不同，提供不同的策略。例如，在一個網際網路拓樸當中，若是能夠有效判斷橋接式/強鍵式連結，那麼就可以針對那些全域橋接式連結做加強，使得遠距離資訊能夠到達目的，就可以提昇整體網路的可靠性。再如傳染病網路的防治，若能針對各類型的連結提供不同的防疫策略，則可以更具成本效益地控制傳染病的傳播。又如在生態網路的研究上，若是全域橋接式連

結兩端的物種瀕臨絕種或是被外來種威脅而數量減少時，那麼很可能會影響到整個生態網路的穩定。因此，藉由我們的演算法的結果，我們能夠提早針對屬於全域橋接式連結兩端的物種加以保護，讓整個生態網路能夠維持穩定的發展與演化。

外，找出網路中每一條連結的特性能夠在網路分群或是尋找網路中的社群 (community) 上提供更多的資訊，全域橋接式連結將會是網路中主要群體間的連結，扮演了主要橋樑的作用。同時，藉由不同階層的橋接式連結，能夠逐漸分割出較小的群體。目前各個科學領域當中都可能存在著複雜網路，藉由本研究所提出的演算法，我們能夠用一致的方法分析不同的網路中的連結型態。更重要的是，隨著分析過的網路資料不斷累積，當日後要研究新的複雜網路時，就能借助先前分析過的網路資料來做比對，若是新的複雜網路與某個已研究過的複雜網路有相似的連結分佈特性，那麼就能藉由研究先前複雜網路的方式來研究新的複雜網路，得到加速的效果。



第二章 文獻探討

依據研究的目標，提出一個利用拓樸性質的演算法能夠判斷出連結在不同網路階層下的性質，首先章節 2.1 探討複雜網路的拓樸性質是如何影響網路的行為，以及前人對於拓樸性質的研究方法為何。章節 2.2 討論強弱連結的概念與不同性質的連結提供網路在運作上何種功能。章節 2.3 藉由前人對於節點的研究導出本研究希望以微觀到宏觀的角度探討連結性質的目的。最後章節 2.4 探討不同計算連結權重方式之間的關係，以及 edge-betweenness 與連結兩端節點共同朋友的差異。藉由本章的文獻探討，除了能夠瞭解前人對於連結特性的研究之外，也能夠呼應本研究的動機和目標。

2.1 拓樸特性影響複雜網路的行為

自從 Erdos-Renyi 隨機模型在 1960 年代被提出之後，使得許多科學家開始思考是否複雜網路的背後只是單純的隨機模型？直觀上來看，複雜網路的拓樸形式必然隱藏了一些重要性質。在從電腦計算能力快速發展之後，對於複雜網路拓樸性質的研究大量被提出，複雜網路中的小世界現象便是近年來廣為探討的性質之一，此現象最重要的三個拓樸特性就是下列三點：

1. 分隔度低

分隔度低表示網路中任兩個節點的距離很短。這邊提到的「距離」代表了兩節點之間最短路徑所經過的連結數。這個性質也就是大家所熟知的小世界現象，即使是規模很大的複雜網路，其中任意兩節點之間卻具有相對較短的路徑數。其中最著名的例子就是 Milgram 的六度分隔實驗，他利用信件追蹤實驗對象的「相識鍊」，最後實驗結果顯示：在美國任兩人之間平均只間隔六個相識的朋友。這樣的性質也在其他大部分的複雜網路中被發現，例如說，全球資訊網、電影演員合作網路、生態網路等。

2. 群聚度高

在自然界中群聚現象是相當常見的現象之一。舉例來說，人際關係網路之中，大部分的人都會與自己的家人或是熟識的朋友生活在一起而形成群體。此外，研究顯示許多複雜網路，節點間都會具有密度較高的連結數而形成群體(Watts & Strogatz, 1998)。

對於群聚度來說，最基礎也是最重要的測量指標就是群聚係數，這個測量指標用來指出網路中節點和其鄰近節點的緊密程度。在以往的研究當中，無論是理論網路或是實際網路，計算群聚係數的方法一直都是相當重要的議題，主要可以分為兩個觀點，全域群聚係數和區域群聚係數。

全域群聚係數為 $\tilde{C} = \frac{\text{三角形個數}}{\text{連通三節點組個數}}$ (Wassermann & Faust, 1994)；連通三節點組(connected triplets of vertices)表示網路中任意三個節點中，其中一節點與另外兩節點相連。此定義方式是以整體網路為觀察對象，計算出目標網路的分隔度。

區域群聚係數是1998年Watts和Strogatz所提出用來決定一個圖(graph)

是否為小世界網路的測量數據之一，其定義如下：網路種某一節點 v_i 的群聚

度 $C_i = \frac{2T(v_i)}{\text{deg}(v_i)(\text{deg}(v_i)-1)}$ ；其中 $T(v_i)$ 為網路中所有通過 v_i 的三角形個數；

$\text{deg}(v_i)$ 為的 v_i 擁有的連結數。整個網路的群聚度為 $\tilde{C} = \frac{1}{n} \sum_{i=1}^n C_i$ ；其中 n 為此

網路的總節點數。此定義方式是以網路中每個節點為觀察對象，以區域角度

計算出整體網路的分隔度。經由研究結果指出，對於大部分的複雜網路來說，

其分隔度相較於其對應的隨機網路的分隔度明顯地高出許多，這點除了說明

分隔度能夠當作複雜網路的指標之外，也說明了複雜網路並不是單純的隨機

網路。

3. 節點度(degree)數呈現冪次率分佈

假設節點度數機率分佈函式為 $p(k)$ ，表示當隨機選取一節點時，其節點度數為 k 的機率；隨機網路的平均節點度數為 $\langle k \rangle$ ，那麼隨機網路的節點度數機率分佈將會出現高峰值 $p(\langle k \rangle)$ 。因為隨機網路中的連結是依照機率值來決定是否出現，因此大部分的節點擁有相近的連結數目，也就是平均值 $\langle k \rangle$ 。然而，有趣的是具有小世界現象的複雜網路的節點度數分佈卻是形成冪次率分佈；尤其是全球資訊網路、網際網路和人際關係網路等...，這些複雜網路均遵守冪次率關係式： $p(k) = p^{-r}$ ，其中 r 是一個常數。此外，這樣的複雜網路也就是所謂的無尺度網路(Goh et al., 2002)。簡單來說，在這些複雜網路當中大部分節點都只有少數的鄰居，其度數較小；但是卻有少部份的節點擁有較多的鄰居，其度數較大。這些度數較大的節點在網際網路中可能是集線器(hubs)，而在人際關係網路中這些節點可能是所謂的“交際花”。

複雜網路的拓樸特性對於複雜網路所展現的功能與行為有很大的影響力。尤其是當複雜網路中個體間需要互相合作解決問題時，拓樸特性更是影響整體效率(performance)的關鍵所在。在此研究中，Dekker 針對拓樸特性與網路行為之間的關係做了深入的研究(Dekker, 2007)。他修改了 Kawachi rewiring process 產生出十種不同的網路形式，接著 Dekker 將這些網路當作下列三個簡易計算模型的拓樸形式：Assignment problem、Kuramoto model 以及 Lagrange's model；並且針對每一個計算模型，提出各自評估解決問題效率高低的數學公式。實驗結果顯示，當網路擁有較低的平均路徑長度時，其解決問題的效率是最高的。另外 Dekker 在 Lagrange's model 的實驗中發現另一個影響網路行為拓樸特性，網路中節點的平均分支度(degree)越高代表節點之間的存取能力越強，相互合作解決問題的效率也就越高。因此，網路中的拓樸特性對於複雜網路的行為有直接的影響。

除了拓撲特性之外，複雜網路的局部拓撲樣式(pattern)也具有關鍵性的影響力。Milo 為了瞭解複雜網路所展現的功能與行為，他想要尋找出複雜網路中具有顯著性的網路基塊(building block)，也就是所謂的基調(motif)(Milo et al., 2002)。Milo 定義出了三節點的基調共有 13 種形式，如圖 3 所示：

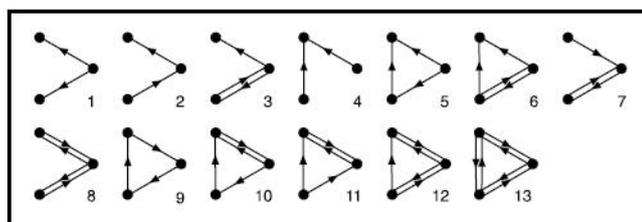


圖 3：13 種大小為三節點的基調(Milo et al., 2002)。

在真實網路中，Milo 紀錄上圖中每一種子圖出現的次數，接著他針對此真實網路所對應的隨機網路(randomized network)紀錄圖 3 中每一種子圖的出現次數，最後若是某一個子圖在真實網路的出現次數比在對應的隨機網路的出現次數明顯高出許多的話，那麼稱此子圖為真實網路的基調。為了要讓實驗更為嚴謹，Milo 在產生對應的隨機網路時，保存每個節點的特性，使得每個節點在各個對應隨機網路當中的 in-degree 和 out-degree 是相同的，如圖 4 所示。

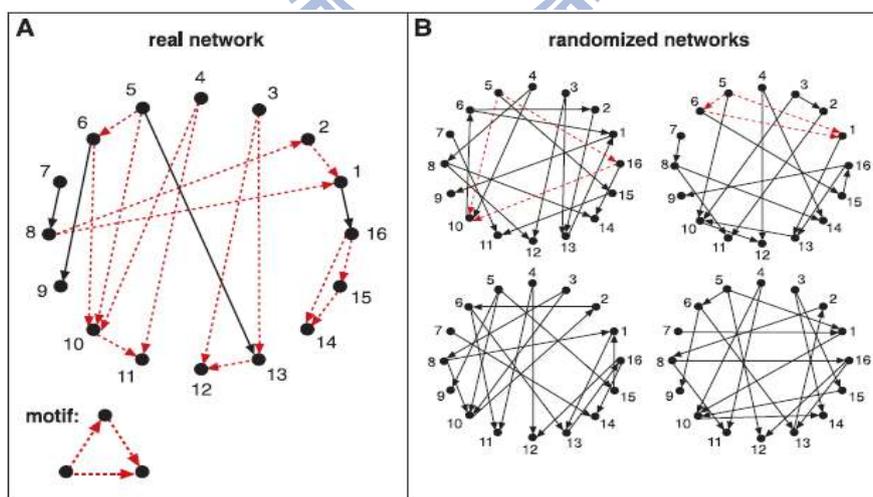


圖 4：產生對應的隨機網路(Milo et al., 2002)。A. 真實網路，B. 對應的隨機網路

此外，Milo 產生的所有對應的隨機網路中，都和真實網路含有相同數目的 $(n-1)$ node 子圖，此虛無假設(null hypothesis)的作法是為了能夠在比較 3-node 的子圖出現次數時，實驗不會受到子結構的影響。實驗結果顯示，許多真實網路

當中都具有特定的基調，舉例來說，基因調控網路與線蟲的神經網路都具有 feed-forward loop 基調，而這兩個複雜網路都是需要將快速的將訊息傳達到目的地，因此可以推測 feed-forward loop 基調對於訊息傳遞的過程具有某種功能性的影響。Milo 的基調實驗除了能夠更深入瞭解拓撲樣式對於整體複雜網路的影響之外，更重要的是能夠利用統計的角度探究複雜網路的演化與設計原則。

Huang 等人針對 Milo 所提出的網路基調加入了強弱連結的概念(Huang et al., 2007)。他們認為當考慮複雜網路的全域特徵與區域結構基調時，連結的權重對於網路的動態行為與功能有著很大的影響。複雜網路的研究學者說明連結的權重代表兩節點互動的能力，並且在真實網路當中，連結的權重能夠直接賦予給連結本身。此外，Milo 所提出的簡單基調雖然具有統計上的顯著性，但是卻可能沒有具備功能性。因此，他們將連結的強弱性質加入基調之後分為兩類：bridge motif 和 brick motif，這樣的觀念不但能夠瞭解全域性質與區域結構之間的關係；並且除了看出那些重要的基調在統計上的顯著性之外，也能夠解釋其在網路上所具有的功能性。

首先，定義出複雜網路中連結強弱的判斷公式(見附錄.A)。接著他們認為基調中若是包涵了至少一條弱連結，那麼此基調就稱作 bridge motif；若是基調中的所有連結都是強連結，那麼此基調就稱作 brick motif。他們發現 bridge motifs 在網路中能夠降低網路的分隔度；brick motifs 在網路中則會建立起群聚的現象。

對於真實網路的研究結果中，他們發現在食物鏈網路中同時存在了 bridge motifs 和 brick motifs。對於 brick motifs 來說，它使得某一種獵食者擁有許多不同種類的獵物，讓他有許多種食物的選擇。當此獵食者消失的時候，整個食物鏈網路會受到很大的影響。而當獵物和獵食者之間的關係被移除之後，此網路影響卻很小，因為獵食者仍然可以經由其他獵物來填飽肚子。另一方面，bridge

motifs 卻提供相反地效果，當獵物和獵食者的關係被移除之後，將會對此食物鏈網路發生重大的影響，可能發生獵食者滅絕或是獵物大量繁殖而發生網路失衡的狀態。另外，他們在電子電路網路當中，只發現了 bridge motifs，而沒有 brick motifs。這是因為在設計電子迴路的時候，工程師總是希望減少多餘的電路，並且希望訊號傳遞是直接的而不是間接的，經過越少節點，則延遲的狀況就會越少，因此這個網路的連結都為弱連結。因此 bridge motifs 提供此網路較少的延遲時間，使得此電路運作時比較有效率。最後他們發現，幾乎所有網路都會存在弱連結，因此都會具有 bridge motifs。但是某些網路的 brick motifs 的數量比 bridge motifs 的數量多出很多。舉例來說，社會網路中，人們都會和自己熟識的人接觸較為頻繁並且互動較多，所以容易造成群聚現象，而原本陌生的兩個人會因為擁有共同熟識的朋友而成為好朋友的機會較高，因此社會網路中的 brick motifs 的數量較多。

藉由 bridge motif 和 brick motif 的研究我們可以看出連結強弱性質對於網路的功能性具有相當大的影響，但是在判斷連結強弱性質時，仍然使用了非強即弱的二分法。因此在我們的研究，希望能夠將連結的強弱性質對於網路的影響力區分成不同的等級，往後在研究複雜網路時，能夠更準確的解釋複雜網路所產生的現象。

2.2 強弱連結

複雜網路中的連結代表了個體之間所具有的關係，個體之間的互動都必須透過連結才能夠達到。然而，許多研究學者都想要知道網路中的連結對於整體網路的影響是否都相同呢？以拓撲的角度來看這個問題時，顯然答案是否定的。舉例來說，兩端的節點分支度(degree)較高的連結可能就會比兩端的節點分支度數較低的連結來說，對於整體網路更有影響力。Granovetter 針對社會網路中的各種連結所扮演的角色做了深入的研究(Granovetter, 1973)，他對於人際關係網路中連結的強弱性質給了初步的概念。舉例來說，人們總是會和家人或是熟識的朋友聚在一起，這種關係視為強連結；而弱連結則是屬於生活中“點頭之交”的相識關係。瞭解了強弱連結的概念之後，Granovetter 想要探討強弱連結何者對於整體網路有關鍵性的影響。如圖 5 所示，假設 A 和 B 與 A 和 C 之間的連結為強連結，那麼 B 和 C 之間存在著強連結的機率將會非常高。也就是說，強連結很少會單獨出現在人際關係網路之中，並且經常會形成兩兩互相認識的三角形關係，也就是所謂的三角閉合原理。在現實生活當中，若 A 和 B 是好朋友並且 A 和 C 也是好朋友的情況下，則 B 和 C 也同樣會是好朋友。這些強連結形成的關係形式會讓訊息滯留在某個區域之內，這也是為什麼我們會時常從自己身邊的好朋友們接收到許多相同重複訊息的原因；反之，強連結對於整體網路的分隔度沒有太大的影響。因此，強連結並不是維持整體複雜網路訊息傳遞的關鍵角色，因為當強連結消失時，訊息仍然可以經由另外剩下的兩條連結傳遞到目的地。

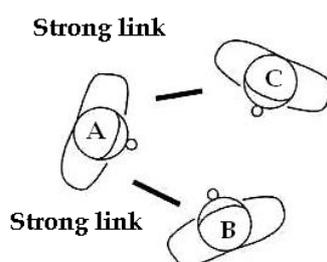


圖 5：社會網路中的強弱連結。

反觀弱連結對於維持整體複雜網路的訊息傳遞扮演了相當重要的角色，若是弱連結消失時，它所連接的兩端節點的訊息將會難以互相傳遞。弱連結就像是「橋」(bridge)的功能，讓原來難以互相到達的兩地能夠藉由橋而達到目的地。Granovetter 的「找工作實驗」將人與人之間的連結強度分為三種：經常見面(一個禮拜至少見一次面)，偶爾見面(一年見面一次以上)以及很少見面(一年至多見一次面)。實驗結果顯示，大約只有 16% 的人找到的工作是藉由經常見面的人所介紹，例如家人或是熟識的好友；而大約有 84% 的人找到的工作是經由偶爾見面或很少見面的人所介紹，例如遠房親戚。此實驗給了我們一個重要的概念：當我們想要傳送某個訊息給身邊的人時，這個訊息通常傳遞的不夠遠，然而，當這個訊息傳遞給那些「點頭之交」時，卻能將訊息傳遞到更遠的地方。Granovetter 最後的結論為：弱連結對於小世界網路能夠擁有低分隔度的性質具有非常關鍵性的因素。弱連結不單單只是複雜網路中個體之間的連結，更是複雜網路中各群體之間的連結。

生化學家 Csermely 針對弱連結對於網路穩定性影響給了概念上的定義 (Csermely, 2006)。他認為某個網路遭到破壞之後，若是此網路還是能夠維持原來的功能或是行為，或是能夠將傷害降到最低的話，那麼此網路的穩定度則越高。反之，此網路的穩定度則越低。當複雜網路遭到破壞的時候，弱連結能夠舒緩 (relaxation) 破壞的情況或是使得網路中的雜訊 (noise) 能夠快速消散，不會滯留在網路當中。因此，弱連結的存在能夠提高複雜網路的穩定性。Onnela 等人的研究結果顯示 (Onnela et al. 2007)，若是網路中的弱連結先被移除，那麼此網路很快就會產生相變並且使得整體網路發生崩潰；而若是先將網路中的強連結移除，此網路則會慢慢收縮卻不會突然發生崩潰的狀況。這證明弱連結對於整體複雜網路穩定性的提昇有關鍵性的影響。

經由上述的研究，似乎強連結的重要性遠不如弱連結。Hansen 以一家大型

電子科技公司中 41 個單位作為複雜網路的研究對象，想要探討單位之間的強弱連結在固定時間內對於新產品的研發所造成的影響(Hansen, 1999)。此研究重點就是加入了單位間相互流動的資訊的複雜程度，並且利用「成文性」(codification)與「獨立性」(independent)的程度來量化資訊的複雜程度。舉例來說，當某資訊的成文性低且獨立性高，那麼此資訊的複雜程度就較高；資訊的成文性高且獨立性低，則此資訊的複雜程度就較低。兩單位之間互相交流的資訊越複雜，那麼就會耗費越多的時間去傳遞。當產品專案小組要發展新產品的時候，就必須要從各單位之間「搜尋」資訊與「傳遞」資訊。當要搜尋資訊的時候，單位之間的弱連結能夠有較高的機會接收到新的資訊或知識，而不會重複搜尋到相同的資訊，可以有效的降低搜尋時間；單位之間的強連結通常需要時間去維持其存在，強連結之間的單位相互交流的機會也比較多，因此當要傳遞資訊的時候，強連結能夠提供複雜程度較高的資訊傳遞，使得資訊能夠完整的抵達目的單位。下圖 6 是將知識的複雜程度和連結的強弱關係圖。



圖 6：連結強弱與資料複雜度的關係圖(Hansen, 1999)。

由研究結果發現，在各個單位之間的連結強度無論是強或是弱，都能夠有效地分享資訊。不過弱連結和強連結在搜尋資訊與傳遞資訊時，各有其優勢與劣勢，最後仍需取決於將在各個單位間傳遞的資訊的複雜程度。強連結提供了網路中各單位之間的高度相關性，因此在傳遞較為複雜的資訊時能夠有較好的效率。當資訊的複雜程度較低的時候，弱連結對於資訊的取得與傳遞能有效的降低產品完成時間。

Onnela 等人利用了手機通訊網路當作研究對象，探討網路結構與連結強弱對於網路所產生的影響(Onnela et al., 2007)。作者認為若是兩個使用者有雙向通話紀錄，代表這兩位使用者在人際網路上可能較為熟識，也許是家人關係或是同事關係。然而，若是兩個使用者之間只有單向通話紀錄的話，他們在人際網路上相識的機會就大為減少。因此為了建構穩固的社會溝通網路，作者認為擁有雙向通話紀錄的兩個使用者之間才有連結。經過上述的方法建構出的網路稱作 mobile call graph (MCG)，擁有 $N = 4.6 \times 10^6$ 個節點， $L = 7.0 \times 10^6$ 個邊。

為了瞭解網路拓撲與連結強度之間的關係是否會影響到全域資訊散播，作者們設計了一連串的實驗步驟去觀察網路中資訊流動的過程(見附錄.B)。研究結果發現：無論弱連結或強連結，都會對於資訊的傳遞造成某些程度上的阻礙。以弱連結為例，在通訊網路中的弱連結會因為本身的權重較低(對話時間花費較少)，使得資訊傳播的量就比較少，而造成傳播上的阻礙。以強連結為例，它屬於同一個社群之內的連結，因此彼此之間比較不容易接收到新的資訊進來。因此，強連結與弱連結在網路中各有其目的與功能，必須要考量切入點為何，才能探討不同特性的連結對於複雜網路是否重要。

關於強弱連結的類似概念也在社會學領域當中出現過。社會學家 Putnam 認為，社會資本是社會生活的重要特性(Putnam, 2000)，例如信任、規範、社會網路，透過協調性的行動，它能夠讓團體內的成員一起行動並且分享目標。公民社

會中的自願性組織的形成與發展，使得社會資本能夠自我補充與儲存的良性循環，也就是信任、規範、社會網路的特質使得自願性組織的參與者能行動一致，並更有效的追求共同的目標與理想，尤其可貢獻其組織之力量，並提昇組織達成目標的效率。

Putnam 將社會資本分成兩個類別：強鍵型社會資本(bonding social capital)與橋接型社會資本(bridging social capital)。強鍵型社會資本指的是要增加排他性的認同，並且維持同質性，其有助於強化個體之間的互助或凝聚力。在社會網路當中，這種黏著性社會資本存在於團體之內，能夠使得原本彼此就認識的人更加靠近結合；對於複雜網路來說，強鍵型社會資本屬於較封閉型的緊密連結。橋接型社會資本指的是要結合社會當中不同層面的人，通常存在著某種程度上的異質性，其有助於連結外界資源並且傳播訊息，並且提供廣泛的認同與互助。在社會網路當中，橋接型社會資本存在於團體之間，使得原本不認識的人形成團體或結合在一起。對於複雜網路來說，橋接型社會資本屬於開放型捐客性連結。對於社會資本更深入的討論請參考附錄.C。

Putnam 提出的兩種社會資本類別的概念與 Granovetter 提出的強弱連結的概念有幾分相似。強鍵型社會資本和強連結是存在於網路中群體之內，而橋接型社會資本和弱連結則是存在於網路中群體之間。然而，Granovetter 強調的是弱連結對於網路中資訊的傳播扮演了重要的角色；Putnam 強調的是不同類別的社會資本提供社會網路何種功能和影響，而兩種社會資本對於社會網路來說是同等的重要。

2.3 從局部觀點到整體觀點

找出網路中心的節點或是節點集合，一直以來都是複雜網路分析的重要研究議題，對於網路中心度的概念也一直被提出(Freeman, 1979)。利用計算網路中各個節點的相對中心度讓我們能夠整合各種中心度的概念，這樣的方法也稱作「節點中心度」。然而，對於節點的中心度也發展出許多不同的測量方法，主要的差別在於局部觀點與整體觀點的不同。若是一個節點為局部中心點，指的是這個節點擁有許多直接相連的節點；若是一個節點為整體中心點，指的是這個節點在此網路的結構上佔據了戰略上重要的位置。舉例來說，如圖 7，雖然 A 點和 C 點的直接鄰居較多，但對於整體結構來說，B 點佔據較重要的位置。因此 A 點和 C 點為局部中心點，B 點為整體中心點。Freeman 提出了局部中心度與整體中心度的測量方式(Freeman, 1979)，對於局部中心度他提出以節點的分支度(degree)為測量基礎，並且為了克服此測量方式因為網路規模不同而造成的局限性，Freeman 定義了節點的局部中心度為節點的分支度與可能聯絡的最多節點數之比。對於整體中心度他提出以不同之間的距離作為測量基礎，計算出各節點之間的接近性(closeness)，若是一個節點在網路中和其他節點的距離都相對較近，那麼則稱這個節點為整體中心點。最後，Freeman 提出了新的中心度概念，稱之為中間度(betweenness) (Freeman, 1979)，若是一個節點居於網路中其他大部分節點之間的最短路徑上，那麼此節點所起到重要的中介作用，因而處在網路的中心。節點中間度的概念被 Girvan 修改成了連結中間度(edge-betweenness)，被用來偵測網路中所存在的社群結構(Girvan & Newman, 2002)。無論是節點中間度或連結中間度，其概念都是在測量節點或是連結的不可取代性，在網路中越是不可取代的節點或是連結則越具有影響力。

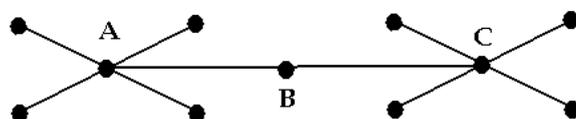


圖 7：整體中心度示意圖。

上個段落的重點是以網路中節點的觀點來評估一個網路的中心。另一方面，我們也可以由整體網路的觀點去測量一個網路在多大程度上具有中心化的結構 (Freeman, 1979)。密度和中心勢的概念能夠以整體的角度去評估網路。密度指的是一個網路的凝聚力有多強；中心勢則是描述這種凝聚力是如何圍繞某些特定的節點而組織起來。因此密度與中心勢是兩個重要且互補的測量指標。但若是一個網路的中心點不只有一個，而是廣泛的分佈在網路當中，那麼網路中心勢的測量可能就不會提供太多資訊給我們。因此，我們必須要觀察一個網路是否存著「結構中心」，所謂結構中心可能是一個節點或是節點的集合，它是網路結構的樞紐。這樣的概念在 Stokman 的研究中已經被勾勒出來 (Stokman, Ziegler & Scott, 1985)，下圖 8 為研究概念示意圖，他的研究重點就是把那些具有最高節點中心度的節點認定為網路的中心，接著再去研究這些節點和網路中其他節點所形成的結構關係為何，最後可以將節點區分成三種不同的類型，不同類型的節點對於網路的影響也有所不同。

網路中心度與中心勢的概念都包含了局部與整體的觀點，也就是說某些節點對於局部網路的影響力較大，但是對於整體網路的影響力就相對較小。然而，無論何種評估方式，都只能計算出區域性質或是直接計算出全域性質。舉例來說，對於節點的局部中心度能夠使用分支度來計算，而整體中心度能夠使用接近性來計算，但是在區域與全域的評估值之間，我們無法知道節點的中心度呈現何種變化。最重要的是已經有許多研究指出複雜網路具有階層性 (Clauset, Moore & Newman, 2008)，我們認為在不同階層之下，連結必然會有不同影響力。因此，針對連結在網路中的重要性，我們希望能夠提出一套測量方式，能夠有效的評估出連結的局部影響力向外擴展至全域影響力。若是能夠區分出複雜網路中不同影響力的連結，那麼在對於複雜網路架構的瞭解與應用上都能夠提供更具有意義的資訊。



圖 8：節點在網路中的位置與中心度關係圖。

2.4 Edge-betweenness 與連結兩端節點的共同朋友

在複雜網路的研究當中，有許多以連結為測量對象的應用，Girvan 提出了複雜網路在拓撲特性上另一個重要的概念：社群結構(Community structure) (Girvan & Newman, 2002)。存在於社群之中的連結密度比較高，節點會緊密的連繫在一起，而社群之間的連結則比較少。在此研究當中，他們提出了一個偵測複雜網路中社群結構的演算法，其中心概念就是計算 edge-betweenness。

對於 betweenness 概念的研究，最早是 Freeman 所提出的 node-betweenness，Girvan 將 betweenness 的概念延伸到連結上，提出 edge-betweenness。計算某個連結 e 的 betweenness 值為：網路中任意兩節點的最短路徑會經過 e 的數目。這個計算方式也就是在評估每一個連結在網路之中所處位置的重要性。我們可以推論得知，在社群之間的連結的 betweenness 值一定相對較高，因為不同社群之間的節點的最短路徑必然會經過這些社群之間的連結，所以這些連結對於社群之間的訊息傳遞擁有較大的作用。

Onnela 等人在 2007 年利用了手機通訊網路當作研究對象，探討網路結構與連結強弱對網路的影響(Onnela, 2007)。在這個研究當中，作者們利用了現今廣為人們所使用的通訊方式：手機，也就是用手機的通訊紀錄來獲取手機用戶之間的互動模式並且建構社會網路的架構。接著作者們利用 edge-betweenness、實際通話時間及流量與連結兩端節點的共同朋友比例(見附錄.D)這三個概念分別計算網路中連結的強度，藉由三種不同的計算方式去瞭解連結強度對於複雜網路的影響。

研究結果展示了三種不同計算方式的關係，我們可以由圖 9 中的藍色圓圈所構成的曲線代表了兩節點真實通話量與共同朋友之間的關係，我們可以發現 $\langle O \rangle_w$ 會隨著 $P_{cum}(w)$ 上升而上升，因為當兩節點的通話量越大則 w 越高，也就表示此兩位手機用戶越熟識，他們之間共同朋友也會越多。圖 9 中黑色菱形所構成的曲線代表了 edge-betweenness 與共同朋友之間的關係，我們可以發現 $\langle O \rangle_w$ 會隨著 $P_{cum}(b)$ 上升而下降，因為當連結的 edge-betweenness 的值越高，也就表示此連結的強度越弱，依據強弱連結理論，這樣的連結屬於社群與社群之間的連結，所以共同朋友也會越來越少， $\langle O \rangle_w$ 的值也就跟著下降了。而圖 9 中紅色方框所構成的曲線代表了隨機給予連結權重的狀況，可以發現 $P_{cum}(w)$ 與 $\langle O \rangle_w$ 無關。

對於 edge-betweenness 來說，是計算連結的不可取代性，越多最短路徑經過的連結對於資訊傳遞就越重要，也就越不可被取代。對於連結兩端節點的共同朋友來說，是計算連結的可取代性，當兩端節點的共同朋友數越多，表示當此連結消失的時候，資訊還是能夠藉由這些共同朋友來傳遞，因此可取代性也就越高。

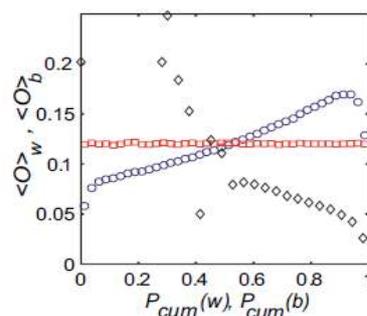


圖 9：共同朋友、edge-betweenness 與實際通話量之間的關係(Onnela, 2007)。 $P_{cum}(w)$ 表示連結權重小於 w 的比例， $\langle O \rangle_w$ 表示在連結權重為 w 下的共同朋友比例。 $P_{cum}(b)$ 表示連結的 edge-betweenness 值小於 b 的比例

第三章 研究方法

3.1 實驗目的與概念描述

在第二章文獻探討中，我們已經從各個角度來探討連結的拓撲特性對於複雜網路具有相當大的影響力，自從 Granovetter 提出強弱連結的概念之後 (Granovetter, 1973)，對於網路中連結強度的研究成為了複雜科學領域的研究焦點之一。為了要更具體的描述連結的強弱概念，我們依據連結在網路中的拓撲性質與其所展現的功能性，本研究提出了「橋接式連結」(bridge link)與「強鍵式連結」(bond link)。複雜網路中橋接式連結能夠將訊息從一個群體傳至另一個的群體，使訊息能夠傳遞到較遠的地方。強鍵式連結能夠將訊息滯留在群體之中，使訊息在群體中不會輕易消失。這兩種類型的連結在網路中具有不同的特性，我們可以從「可取代性」的角度來分析連結。當橋接式連結消失時，網路中的訊息很可能無法到達目的地，或者需要花更長的時間完成傳遞；當強鍵式連結消失時，訊息仍然可以藉由周遭的連結來完成傳遞。換句話說，橋接式連結在網路中的可取代性較低，而強鍵式連結的可取代性較高。直觀來看，網路中的連結若是越緊密，當某條連結消失時，訊息能夠從其他連結傳遞的機會就越高，此連結的可取代性也就越高。

在這篇研究論當中，我們利用「共同朋友」的概念來探討連結的特性，當連結兩端節點的共同朋友數越少，它們可能存在於網路中不同的兩個群體之中，那麼此連結在網路中可能扮演了群體之間訊息橋樑的角色；兩端節點的共同朋友數較多，它們可能存在於網路中同一群體之間，那麼此連結可能是群體之中較為緊密的連結。除此之外，為了要避免對於連結特性只有強和弱的二分法，本研究提出網路「階層」的概念，我們能夠探討連結在不同階層時，其所具有的特性為何，連結的影響力範圍有多廣，進而提出了「全域橋接式連結」(global bridge link)、
「N 階橋接式連結」(N degree of bridge link)與「沉沒橋接式連結」(sink bridge

link)，並且根據不同的連結類型，我們利用「可取代性」給予這些連結類型概念上的描述：

1. 強鍵式連結：若是連結為強鍵式連結，若且唯若當此連結被移除時，此連結的兩端節點能夠經由路徑長度為 $O(1)$ 的替代路徑來傳遞訊息。
2. N階橋接式連結：若是連結為N階橋接式連結，若且唯若當此連結被移除時，此連結的兩端節點能夠經由路徑長度為 $O(m)$ 的替代路徑來傳遞訊息(m 為此連結所連接的兩區域群體內的節點個數總和)。
3. 全域橋接式連結：若是連結為全域橋接式連結，若且唯若當此連結被移除時，此連結的兩端節點能夠經由路徑長度為 $O(n)$ 的替代路徑來傳遞訊息或者無法找到替代路徑(n 整體網路的節點個數)。
4. 沉沒橋接式連結：若是連結為沉沒橋接式連結，若且唯若當此連結被移除時，使得此連結兩端節點中的其中一節點的分支度為 0。

經由瞭解不同連結類型的概念，我們可以從可取代性的角度瞭解這些連結類型在網路當中所具有的影響力。當某連結被移除後，其替代路徑的路徑長度越長，表示資訊必須花費較久的時間才能將訊息傳遞到目的地，此連結對於整體網路的影響力也就越大。

利用上述的概念，在接下來的子章節中我們先定義出網路的「階層」(hierarchy)，其目的是為了觀察連結在不同階層下所具有的性質為何。接著我們利用階層的概念定義出不同階層下的共同朋友比例，其目的是為了量化連結在不同階層下所具有的影響力，舉例來說，當連結在某階層下的共同朋友比例越高，則當此連接被移除之後，在此階層能夠找到路徑長度較短的替代路徑，那麼連結在此階層所具有的影響力也較小。

3.1.1 連結的「階層」(Hierarchy)

首先，我們利用圖論的定義將複雜網路表示成一個無向圖 $G = (V, E)$ ，其中 V 為節點的集合，若 G 有 n 個節點，則 $V = \{v_1, v_2, v_3, \dots, v_n\}$ 。 E 為連結的集合，若 G 有 m 個連結，則 $E = \{e_1, e_2, e_3, \dots, e_m\}$ 且 e 為二元組數對， $e = (i, j)$ ，其中 $i, j \in V$ 。

將複雜網路定義成圖之後，我們希望能夠從區域觀點出發，接著逐漸拓展成全域觀點，一步一步瞭解每一條連結對於整體網路的影響為何，因此針對不同的尺度觀點，我們必須定義「階層」(hierarchy)。假設圖 $G = (V, E)$ ， G 中含有某一連結 $e = (i, j)$ 。 G 中節點 i 或節點 j 距離為 k 的所有節點之集合所形成的子圖 $g_{e,k} = (V'_{e,k}, E'_{e,k})$ 稱作連結 e 的第 k 階層。

$$V'_{e,k} = \{v | (v \in V) \wedge (\text{distance}(i, v) \leq k \vee \text{distance}(j, v) \leq k)\} \text{ 且 } 1 < k < D。$$

$$E'_{e,k} = \{(a, b) | (a, b) = V'_{e,k} \times V'_{e,k}, (a \neq b), (a, b) \neq (i, j)\}。$$

上式中的 $\text{distance}(i, v)$ 表示節點 i 與節點 v 的最短路徑長度。 D 表示此網路的網路直徑，本研究的網路直徑¹定義為：任意兩節點之間所存在最短路徑的平均值。

$g_{e,k}$ 相當於連結 e 向外延伸 k 個路徑之後所形成的子圖，而 k 值會被限制在 1 到 D 之間，對於連結 e 我們會得到 $g_{e,1}, g_{e,2}, \dots, g_{e,D}$ 總共 D 個子圖。藉由這些子圖，我們可以分析連結 e 從區域到全域的影響力。舉例來說，在複雜網路圖中我們定義某連結 $e = (i, j)$ 的第 n 階層表示由 e 為中心向外延伸 n 步所形成的子圖。圖 10.a 為一個簡單的網路圖，白色的節點是由連結 e 為中心向外延伸一步所能到達的節點，因此對於連結 e 來說，其第一階層就是白色節點所形成的子圖(圖 10.c)。而白色節點和黃色節點所形成的子圖則是第二階層(圖 10.c)。

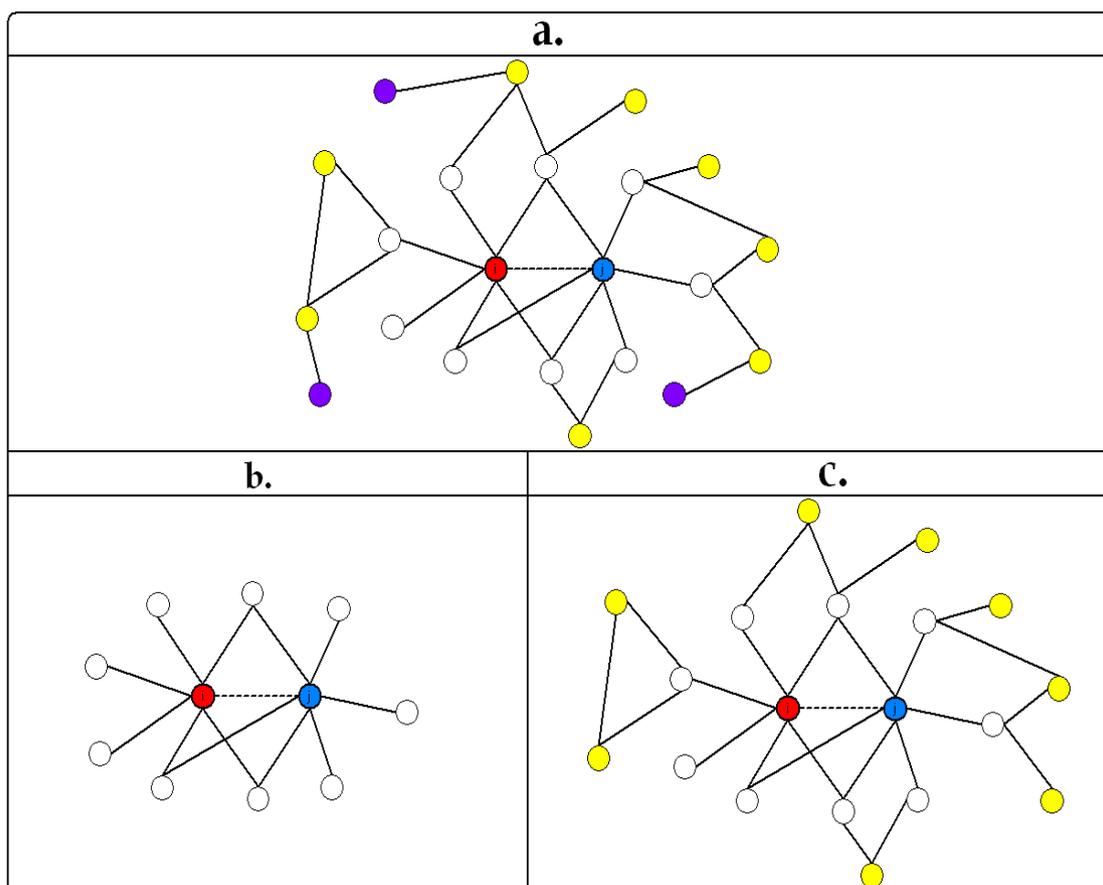


圖 10：階層概念圖。a. 整體網路圖，b. 第一階層，c. 第二階層。

值得注意的是根據階層的定義，對於目標連結作延伸時，我們必須忽略目標連結，這個目的就是為了要計算目標連結的可取代性。所謂可取代性就是想要評估當目標連結消失之後，其功能是否會被網路中其他的連結所取代。所以我們對於目標連結作延伸時不考慮任何經由目標連結所延伸到的節點。

依據網路直徑的定義，每一條連結在做延伸時可能不會延伸到所有的節點，但是依據複雜網路的重要拓樸特性—低分隔度。假設網路直徑為 D ，對於任一節點延伸 D 個路徑長之後所產生的子圖，均能夠涵蓋大部分的節點，對於連結的影響範圍計算已經足夠。直觀來看，若是節點 i 和節點 j 的路徑距離相對較遠，那麼節點 i 附近的連結對於節點 j 的影響則是相當有限。

3.1.2 連結兩端節點的共同朋友

複雜網路中某一連結 $e = (i, j)$ ， i_n 表示節點 i 延伸 n 步之後所有拜訪到的節點之集合。假設網路圖 $G = (V, E)$ ，我們可以將 i_n 定義成：

$$i_n = \{v \mid (v \in V) \wedge \text{distanse}(i, v) = n\} \circ$$

舉例來說，如圖 11 所示，橘色節點是節點 i 延伸一步之後所能拜訪到的節點。

因此， i_1 為橘色節點的集合。

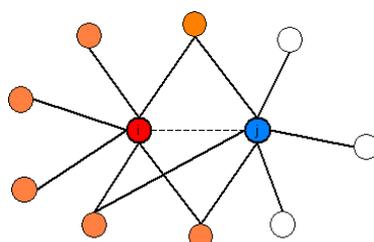


圖 11：節點 i 延伸一步之後所拜訪到的節點形成 i_1 。

$i_n \cap j_n$ 代表節點 i 延伸 n 步與節點 j 延伸 n 步之後的共同朋友所形成的集合。

如圖 12 所示，a 圖代表節點 i 延伸一步之後所拜訪到的節點(i_1)，b 圖代表節點 j 延伸一步之後所拜訪到的節點(j_1)，而 c 圖中的綠色節點代表了連結 e 在第一階層的共同朋友所形成的集合($i_1 \cap j_1$)。

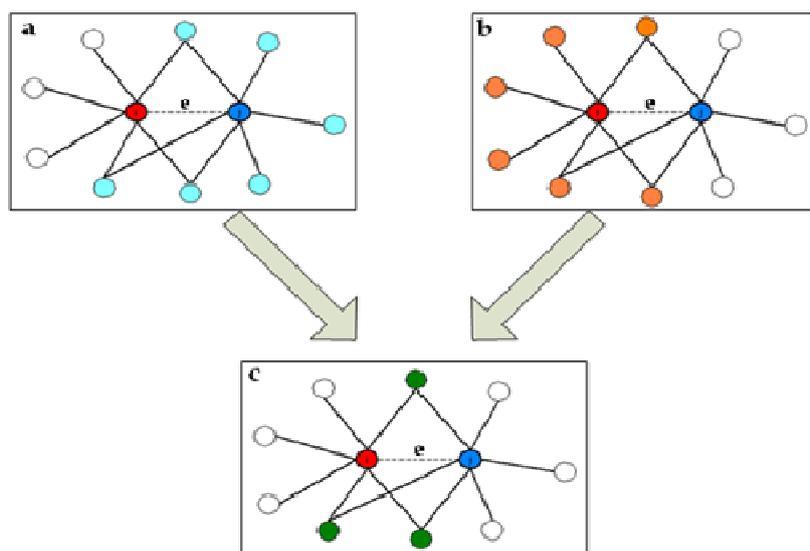


圖 12：共同朋友。

由上述的定義，對於複雜網路中每一條連結 $e = (i, j)$ 我們可以推導出連結 e 在第 n 階層下「所有」共同朋友的集合 $C_{i,j,n}$ ：

$$|C_{i,j,n}| = |i_{n-1} \cap j_n| + |i_n \cap j_{n-1}| + |i_n \cap j_n|$$

因此，複雜網路中每一條連結 $e = (i, j)$ 在第 n 階層的共同朋友比例定義如下：

$$\text{共同朋友比例 } R_{e,n} = \frac{|C_{i,j,n}|}{\min(|i_{n-1}|, |j_n|) + \min(|i_n|, |j_{n-1}|) + \min(|i_n|, |j_n|)}$$

圖 13 為在每一階層共同朋友的意識圖。左邊三角形的每一階層代表連結 e 的左節點所延伸到的節點之集合。舉例來說，黃色部份是左節點延伸一步所到達的節點形成的集合。同理，右節點每一階層延伸所形成的集合為右邊的三角形。而兩個三角形在每一階層重疊的部份就是在此階層中兩節點的共同朋友所形成之集合。

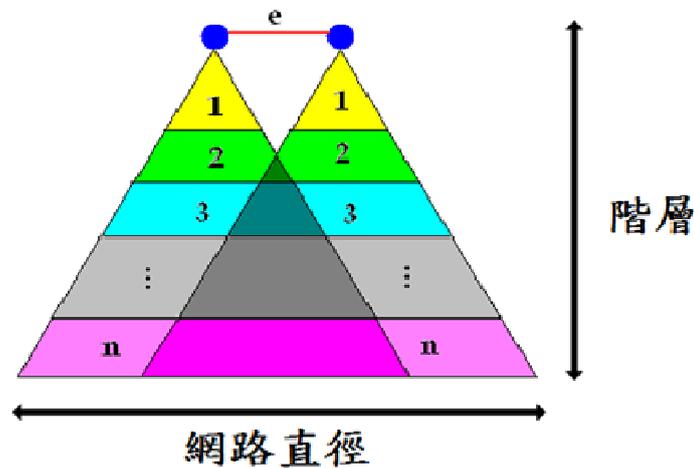


圖 13：共同朋友概念圖。

1. 本研究所採用的網路直徑定義為：網路中任意兩節點之間最短路徑的平均值。如果網路為非連通圖，那麼網路直徑為無限大；若網路為完全圖，則網路直徑為 1。有些複雜網路的研究將網路直徑定義為任意兩節點之間最短路徑中的最大值，這樣的作法如果遇到複雜網路當中只有極少部份的節點距離較遠，那麼對於大部分連結則會執行無意義的延伸動作而沒有效率。

3.2 實驗設計

目標網路中連結兩端節點的共同朋友比例越高時，當此連接消失，兩端節點的訊息仍然能夠藉由這些共同朋友當作「替代」途徑達到目的地，所以此連結消失後對於網路的影響就越小。反之，目標網路中兩端節點的共同朋友比例越低，當此連結消失之後，兩端節點難以找到這些「替代」途徑來幫忙傳遞訊息，可能導致整體網路無法達到某種功能或是目的，造成較大的傷害。為了要判斷連結兩端節點的共同朋友比例的高低，我們利用了目標網路本身和其「對應的隨機網路」產生判斷連結共同朋友比例高低的「內部門檻值」與「外部門檻值」。圖 14 為實驗操作流程圖：

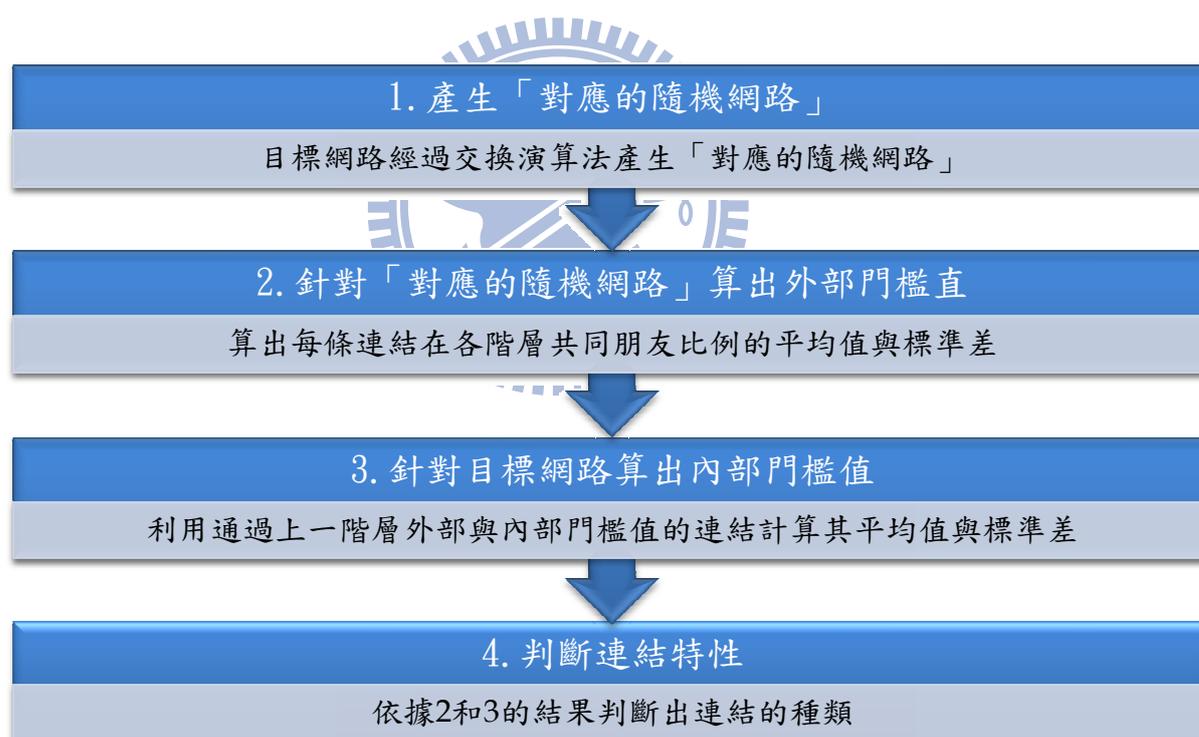


圖 14：實驗流程圖。

3.2.1 產生「對應的隨機網路」

我們使用交換演算法(i.e. 若 $a \rightarrow b$, $c \rightarrow d$ 則變為 $a \rightarrow d$, $c \rightarrow b$)來產生目標網路的「對應的隨機網路」(Milo et al., 2002; Huang et al., 2007)。產生這些隨機網

路的目的是為了觀察目標網路中的拓樸特性(共同朋友比例分佈)是否會和其對應的隨機網路相同，若是目標網路的拓樸特性與其對應的隨機網路有相當明顯的差距，那麼這個拓樸特性則具有統計上得顯著性。因此對於每一個目標網路我們都會產生至少 100 組對應的隨機網路，以降低計算上的誤差。演算法所產生的隨機網路中，節點數和連結數都會和目標網路相同；此外，產生的隨機網路的節點分支度(degree)也會和目標網路相同。因此，對應隨機網路的分支度序列和目標網路完全相同。經過交換演算法所產生的隨機網路和目標網路在結構上具有相似性，因此能夠避免結構上的差異所造成的影響。如圖 15 所示，左圖連結(A,C)和連結(B,D)經過交換演算法之後，成為右圖連結(A,D)和連結(B,C)。

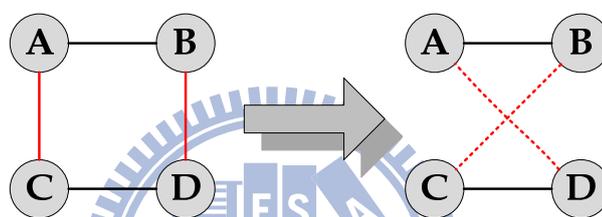


圖 15：交換演算法示意圖。

將目標網路輸入交換演算法之後，隨機選取連結的次數和目標網路的連結數相同，而曾經被選取到的連結還是有可能再次被選取到，其目的都是為了在產生對應的隨機網路的過程中，其隨機化的程度更高。

3.2.2 判斷連結種類

產生完對應的隨機網路之後，分別針對目標網路和其對應的隨機網路計算共同朋友比例值與連結種類的判斷門檻值的相關參數。

針對「對應的隨機網路」，計算出外部門檻值：

1. 計算出所有「對應的隨機網路」的平均網路直徑 D 。
2. 計算出網路中連結從第 1 階層到第 D 階層的共同朋友比例。
3. 計算出所有「對應的隨機網路」的第 1 階層到第 D 階層的共同朋友比例的平均值($k_1 \sim k_D$)與平均標準差($\sigma_1 \sim \sigma_D$)。

針對目標網路，計算出內部門檻值：

1. 計算目標網路中的連結從第 1 階層到第 D 階層的共同朋友比例值。
2. 找出目標網路中每一階層中通過上一階層外部與內部門檻值的連結，計算這些連結的共同朋友比例的平均值($S_2 \sim S_D$)與標準差($\delta_2 \sim \delta_D$)。在第一階層中沒有上一階層，所以 S_1 和 δ_1 是直接由通過「一階外部門檻值」的連結所計算得到。

計算出所有相關數據之後，接下來依據統計的概念，我們藉由下面的步驟來判斷複雜網路中每一條連結屬於何種類型。

1. 對於目標網路中的連結 $e = (i, j)$ 延伸一步至 L 階層(L 從 1 開始)。若是此連結的兩端節點中，有一個節點無法延伸(i_L 或 j_L 為空集合)，則此連結被判斷為「沉沒橋接式連結」(sink bridge link)。若兩端節點皆可延伸，則繼續至步驟二。
2. 計算連結 e 在第 L 階層的共同朋友比例 $R_{L,e}$ 。若是 $R_{L,e} \geq k_L + 2\sigma_L$ ，則連結 e 被判斷為「強鍵式連結」(bond link)，判斷流程結束。若是 $R_{L,e} < k_L + 2\sigma_L$ ，則繼續至步驟三。($k_L + 2\sigma_L$ 稱之為「 L 階外部門檻值¹」)
3. 若是 $R_{L,e} \geq S_L - \delta_L$ ，則連結 e 被判斷為「 L 階橋接式連結」(L degree of bridge link)並且繼續將連結 e 延伸至 D 階層，並且在每階層檢驗連結 e 是否轉變為強鍵式連結。若是 $R_{L,e} < S_L - \delta_L$ ，則直接繼續至步驟四。($S_L - \delta_L$ 稱之為「 L 階內部門檻值²」)
4. 若是 $L = D$ ，則連結 e 被判斷為「全域橋接式連結」(global bridge link)，判斷流程結束。否則回到步驟一。

下圖 16 是完整演算法的流程圖(包含圖 14 的實驗步驟)。

1. 「 L 階外部門檻值」主要是判斷連結的共同朋友比例值是否高於所有對應的隨機網路的 L 階層下所有連結的共同朋友比例值之平均加上 x 倍標準差(x 通常為 2)，是判斷強鍵式連結或橋接式連結的標準。

2. 「 L 階內部門檻值」主要是判斷連結的共同朋友比例值是否高於目標網路的 L 階層下所有連結的共同朋友比例值減掉 y 倍標準差(y 通常為 1)，是判斷 L 階橋接式連結的標準。

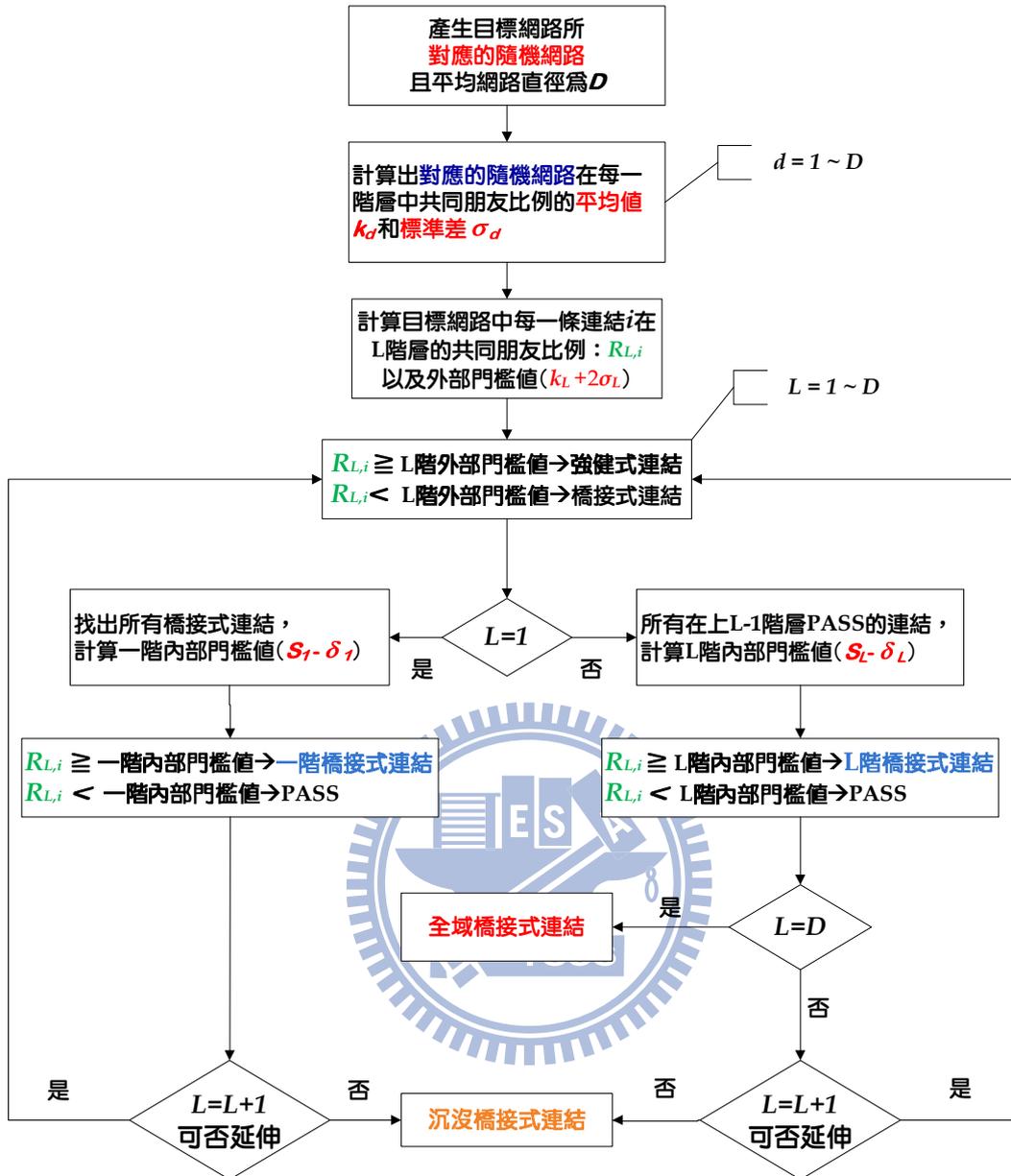


圖 16：演算法流程圖。

在上述的流程我們利用了真實目標網路與其對應的隨機網路的統計參數來判斷連結的種類。在前面的章節我們提過複雜網路中，某些連結的特性可能會影響整體網路，而有些只會具有區域的影響力。因此整個判斷流程的精神為：在每一階層利用統計顯著性判斷出連結的影響力是否能夠影響到下一階層。

在第一步驟中，針對某連結的兩端節點作延伸一步，若是無法延伸代表兩節點中至少有一個節點為末端節點，當此連結消失的時候，則末端節點則與其他節

點完全分離，雖然這樣的連結對於整體網路的影響不大，但是對於末端節點卻有非常重要的地位，因為此連結是末端節點與整體網路唯一的聯繫，所以我們將此連結判斷為「沉沒橋接式連結」。

在第二步驟中，我們利用對應隨機網路第 L 階層中所有連結的共同朋友比例的平均值加上兩個標準差的結果當作真實目標網路中連結的「 L 階外部門檻值」，其目的是要判斷出「強鍵式連結」，若是某個連結在第 L 階層的共同朋友比例大於這個值，表示這條連結的兩端節點在第 L 階層具有相對較多共同朋友數，在拓樸上與附近的連結也較為緊密，當此連結消失的時候，其兩端節點的資訊也能夠經由這些共同朋友來傳遞，所以此連結的性質對於網路的影響較小；以統計的角度來看，若是某連結的共同朋友比例超過門檻值代表了這樣的連結出現在對應的隨機網路中的機率只有 2.5% 左右，說明了在拓樸中這樣的連結具有統計上的顯著特性。

在第三步驟中，我們利用真實目標網路在第 L 階層中通過「 $L-1$ 階內部門檻值」與「 $L-1$ 階內部門檻值」的連結的共同朋友比例的平均值減掉一個標準差的結果當作真實目標網路中連結的「 L 階內部門檻值」，其目的是為了要判斷出「 L 階橋接式連結」。若是某個連結在第 L 階層的共同朋友比例小於這個值，表示此連結在第 L 階層具有的共同朋友比例相對較少，當此連結消失的時候，訊息可能需要經更多替代路徑來傳遞或甚至是無法傳遞到目的地，因此這樣連結的性質對於網路的影響力較大。而共同朋友比例大於門檻值的連結，其影響力只侷限在 L 階層以下，因此被判斷為「 L 階層橋接式連結」。此階段為了要避免對應的隨機網路為均勻的性質，無法看出在拓樸上連結對於區域的影響，並且避免真實網路中強鍵式連結的共同朋友比例值對於計算內部門檻值的影響，因此我們直接利用真實目標網路中通過「 $L-1$ 階內部門檻值」與「 $L-1$ 階內部門檻值」的連結的共同朋友比例平均值減掉一個標準差的結果當作「 L 階內部門檻值」。

在第四步驟中，若是 L 與 D （「對應的隨機網路」的平均網路直徑）相等時，此連結的延伸範圍可能已經涵蓋整個目標網路，表示在拓樸上此連結對於整體網路有關鍵性的影響，對於整體網路的資訊傳遞可能佔有重要的地位。因此，我們將這種連結判斷成「全域橋接式連結」。

依據上述的判斷流程，我們能夠依照目標網路中每一條連結的拓樸特性，將其分類成：橋接式連結、 N 階橋接式連結、沉沒橋接式連結與強鍵式連結。依據 Granovetter 的強弱連結理論 (Granovetter, 1973)，弱連結重要性在於訊息的傳遞，也就是說弱連結消失的時候，網路中群體之間的訊息將難以交流。因此這樣的概念就可以說明弱連結的特性和橋接式連結的特性是相當類似的；強連結的特性和強鍵式連結的特性是相當類似的。

上述的判斷流程當中，我們可以注意到當連結被判斷成強鍵式連結時，對於此連結的判斷才會停止，否則一定要等到延伸路徑數與對應的複雜網路的平均網路直徑相等時才停止。這樣設計演算法的原因是隨著目標連結延伸的路徑數增加，所形成的子圖也就越來越大，此連結兩端的共同朋友就會越來越多，則此連結與其他連結也就越趨緊密。因此，當某個連結在較低階層被判斷為強鍵式連結之後，那麼在較高階層當中，此連結仍然也會是強鍵式連結；當某個連結在第 n 階層被判斷為 n 階橋階式連結之後，那麼在第 N 階層當中 ($N > n$)，此連結有可能被判斷為強鍵式連結並且依據上述原因而不可能被判斷為 N 階橋接式連結。依據強弱連結的概念，在擴大網路的情況下，連結的特性只會由弱轉強，不會由強轉弱。因此，本研究提出的演算法不單單只是依據連結在不同階層時所具有的性質分成不同的類型，更能夠觀察到連結從低階層(區域)到高階層(全域)的性質變化，讓我們可以更深入的瞭解每一條連結對於網路的影響。

3.2.3 討論

利用前面章節提出的演算法能夠偵測出網路中每一條連結在每一個階層下的「可取代性」。若是連結兩端節點的共同朋友比例越高，則此連結的可取代性就越高；若是連結兩端節點的共同朋友比例越低，則此連結的可取代性就越低。本研究所提出的演算法純粹以拓樸與統計的角度判斷連結的類型，簡化複雜網路中個體之間的關係方向性與背景特性(domain knowledge)，將分析焦點放在拓樸架構如何影響連結的特性，並且依據統計性質將連結分類成不同的類型。此外，利用拓樸架構的概念能夠避免不同複雜網路中的連結量化方式，能夠以一致的方法分析不同的網路，並且瞭解各種連結類型在不同複雜網路中所佔有的比例以及成因，最重要的是能夠觀察不同複雜網路之間連結類型所佔比例的差異，進而提供未來複雜網路研究更進一步的拓樸資訊。

但是純粹以拓樸與統計的角度來分析複雜網路可能碰到的問題是利用演算法判斷出的連結類型在網路中可能不具有功能上的意義。舉例來說，網路中某連結被判斷成全域橋接式連結，代表此連結以拓樸架構的觀點來說，其影響力是整體網路中最大的，但是在真實世界中此連結本身的性質對於整體網路的影響可能不是最大的，甚至是沒有影響力的。

第四章 研究結果與發現

在以往的研究當中，連結的性質一直是研究複雜網路動態行為的焦點，不同性質的連結所具有的功能性以及對於網路的影響性都不同。雖然每一種領域的複雜網路都具有各自量化連結性質的方式，但是我們仍然能夠藉由拓樸架構來探討不同複雜網路之間的關係。本研究所提出的演算法能夠利用複雜網路的拓樸架構與統計方法，將連結區分成強鍵式連結(Bond link)、全域橋接式連結(Global bridge link)、N 階橋接式連結(N degree of bridge link)與沉沒橋接式連結(Sink bridge link)。強鍵式連結的兩端節點共同朋友較多，可取代性高，容易形成群聚現象而存在於群體之內；全域橋接式連結的兩端節點共同朋友較少，可取代性低，能夠降低網路的分隔度，通常是群體之間的溝通橋樑；N 階橋接式連結的可取代性介於強鍵式連結與橋接式連結之間。本研究希望藉由演算法去偵測各種不同複雜網路中所存在的各種類型的連結，再利用連結不同類型的連結所具有的特性來分析複雜網路可能具有的性質與行為。

4.1 理論網路

本章節首先針對已知性質的理論網路為研究對象，利用本研究提出的演算法判斷出網路中每一條連結的種類。主要目的除了對於本研究的方法做初步的驗證之外，也探討了不同性質的連結分佈與其他複雜網路的全域拓樸性質(分隔度與群聚度)之間存在著什麼樣的關係。

4.1.1 WS-小世界網路

利用理論網路當作目標網路，驗證目標網路中已知的拓樸性質與強鍵式連結和橋接式連結之間的關係是否合理以及觀察這些特性改變的趨勢。因此，我們將N 階橋接式連結與全域橋接式連結均視為橋接式連結。在這個驗證中我們專注於

沉沒橋接式連結、強鍵式連結與橋接式連結在各種不同狀況下的理論網路中所扮演的角色為何。

我們利用 Watts 和 Strogatz 所提出產生小世界網路的方法(Watts & Strogatz, 1998)，將一個節點個數為 100、平均分支度為 $\langle k \rangle$ 的規則網路，使用不同的繞線機率 p 產生一連串的小世界網路($p = 1\%、2\%、4\%、8\%、16\%、32\%、64\%$)。

1. 節點個數：100、連結個數：200、平均分隔度 $\langle k \rangle$ ：4。

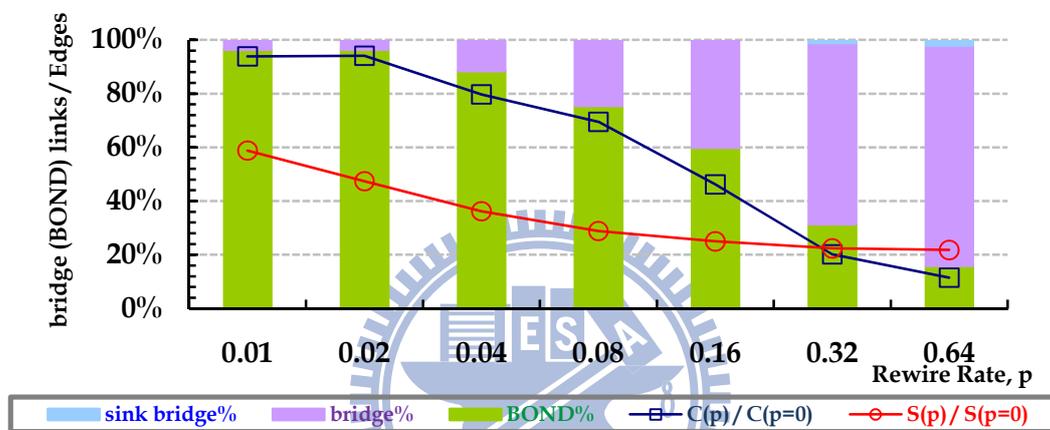


圖 17：WS-Small world, $N=100, \langle k \rangle=4$ 。

2. 節點個數：100、連結個數：400、平均分隔度 $\langle k \rangle$ ：8。

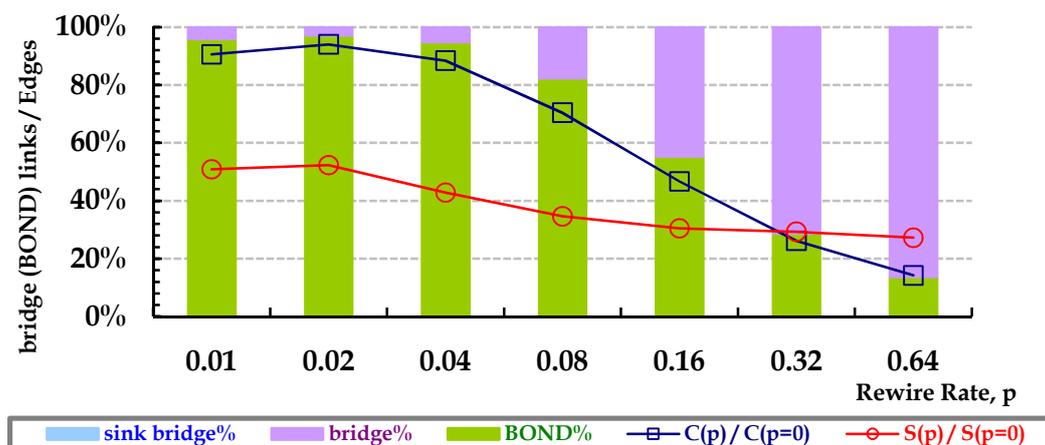


圖 18：WS-Small world, $N=100, \langle k \rangle=8$ 。

3. 節點個數：100、連結個數：800、平均分隔度 $\langle k \rangle$ ：16。

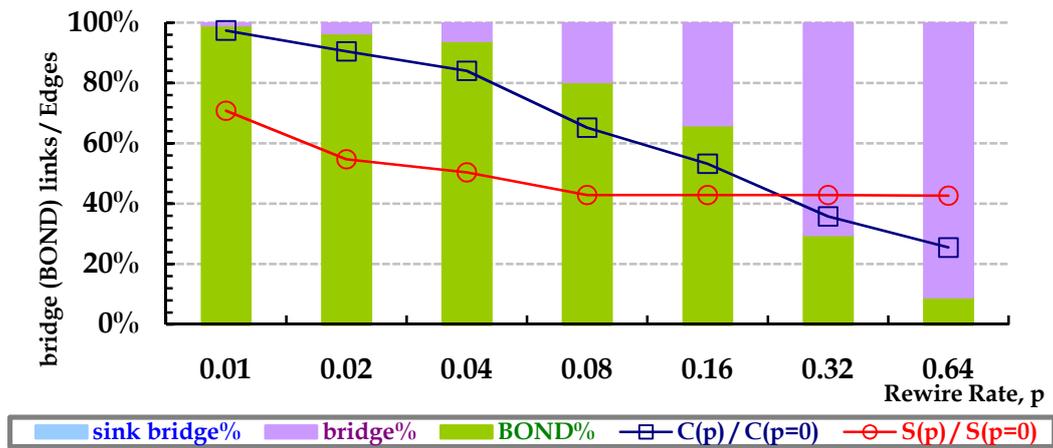


圖 19：WS-Small world, $N=100$, $\langle k \rangle=16$ 。

4. 節點個數：100、連結個數：1600、平均分隔度 $\langle k \rangle$ ：32。

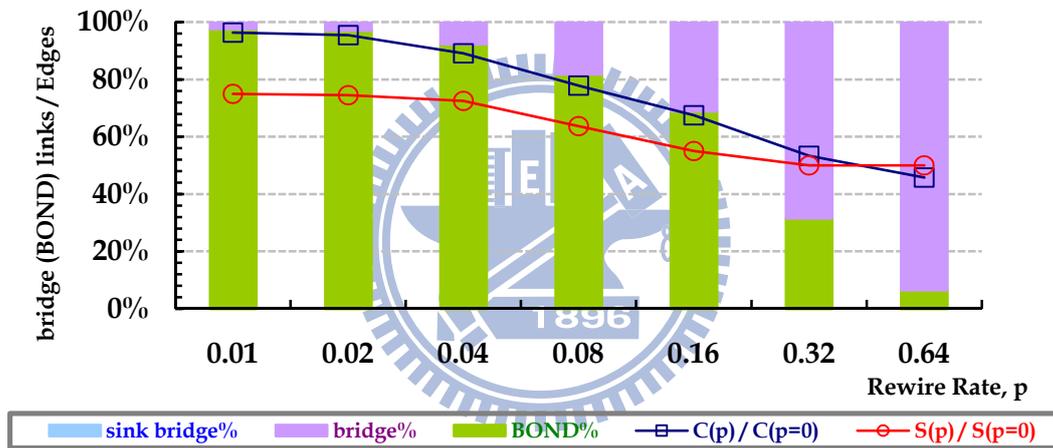


圖 20：WS-Small world, $N=100$, $\langle k \rangle=32$ 。

5. 節點個數：100、連結個數：3200、平均分隔度 $\langle k \rangle$ ：64。

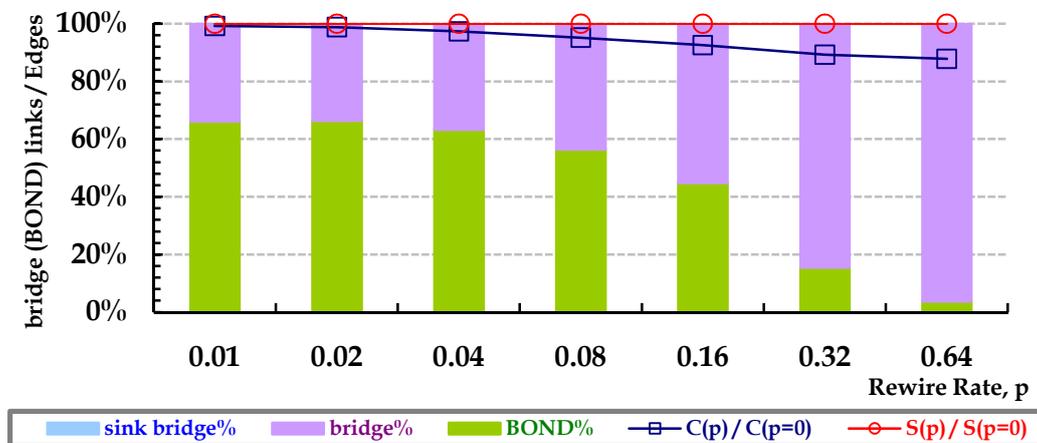


圖 21：WS-Small world, $N=100$, $\langle k \rangle=64$ 。

由上圖 17 的結果所示，當繞線機率只有 1%或是 2%時，其網路仍然屬於規則網路，只有少數的連結經由繞線成為橋接式連結。當繞線機率越來越大時，整個網路也越來越趨向隨機網路時，橋接式連結所佔有的比率越來越高，相對的強鍵式連結就越來越少，因此橋接式連結與強鍵式連結才會呈現彼此消長的現象。圖 18 到圖 21 也都呈現相同的結果，當繞線率的增加時，橋接式連結所佔有的比例也隨之增加，相對的強鍵式連結就跟著減少。關於實驗結果統計圖的相關數值請參考附錄 E 的表 E.1。

這個結果說明了當繞線機率低的時候，網路中連結兩端節點的共同朋友比例呈現均勻的狀態，每一個節點都和附近固定的節點數相連，因此大部分的連結都是強鍵式連結。然而隨著繞線機率上升，被重新繞線的連結產生了網路中的「捷徑」，這些捷徑兩端節點的共同朋友和之前完全不同，降低了共同朋友比例，因此橋接式連結也就逐漸增多。隨著橋接式連結和強鍵式連結彼此的消長，網路的分隔度(紅色圓形的線段)和群聚度(藍色方形的線段)也都呈現下降的趨勢，因為橋接式連結扮演了那些捷徑的角色，任意兩節點之間的最短路徑能夠利用橋接式連結縮短路徑長度，在之前的研究當中也指出這些捷徑的存在會降低複雜網路的分隔度。同理，隨著強鍵式連結的減少，使得破壞了原本緊密相連的節點，使得群聚度也跟著下降。

上述的驗證步驟是說明連結性質與分隔度和群聚度的相關性。不過隨著節點分支度 $\langle k \rangle$ 的上升，分隔度和群聚度的下降趨勢都呈現緩和的情況，是因為網路節點之間的連結越來越多。我們觀察圖 21，節點的平均分支度為 64，網路本身的分隔度很低且群聚度很高，但是隨著繞線機率 p 的增加，其所產生的捷徑增加，我們的演算法仍就判斷出網路中有許多橋接式連結被產生。這也說明了我們所提出的橋接式與強鍵式連結所具有的特性並不會依據分隔度或群聚度的高低而做出相對應的變化，能夠提供研究者更多的拓樸資訊。我們也試著固定繞線機率 p ，

觀察當節點分支度變化時，橋接式連結與強鍵式連結的分佈情況(請參照附錄 E)，其結果也和顯示了相同的結論。

Watts 和 Strogatz 的方法所產生的小世界網路有可能會破壞網路的連通性，因此我們也利用 Newman 和 Watts 提出建立小世界網路的方法，產生出一連串的小世界網路輸入本研究的演算法中，也產生相同的結果(請參照附錄 E)。經過針對理論網路的驗證實驗，說明了本演算法所判斷出來的連結種類是具有理論基礎並且符合實際情況，也證明本演算法的正確性。



4.2 真實世界網路

本研究認為不同領域的複雜網路，會依據拓樸架構的差異，使得各種類型的連結分佈也會有很明顯的不同，也因此會影響複雜網路所展現的動態行為。這個章節我們針對三種不同領域的複雜網路當作實驗對象：社會網路、生態網路與電子電路網路，利用演算法找出各種類型的連結所佔有的比例，觀察並且分析不同種類的連結對於複雜網路所可能具有的影響力。

4.2.1 社會網路(朋友網路)

在社會網路的網路圖當中，節點代表社會中的人，若兩節點之間有連結代表兩人之間有某種關係。第一個社會網路是美國某成衣工廠內員工之間在閒暇時刻的遊戲互動網路。第二個社會網路是1992年美國國家自然科學基金會夏令營中的14位參與者與4名指導教師之間的互動情形。第三個社會網路是美國南部某村落婦女互動網路。第四個社會網路是監獄中的囚犯所組成的關係網路，每一位囚犯被詢問「在自己的監獄分區當中，哪些人是你認為親近的朋友」。(拓樸性質見表1)

圖22是四種社會網路經過判斷演算法之後的結果統計圖，在囚犯網路當中，囚犯的活動範圍是相當受到限制的，很容易和自己的室友或鄰近的獄友成為朋友，共同朋友的比例較高，因此有接近一半的連結比例是強鍵式連結。而吃飯或是放風的時間，囚犯們就能夠擁有較大的活動空間，也能夠較自由的與較遠牢房的囚犯進行互動，因此才會出現三階橋接式連結或甚至是全域橋接式連結。而某些囚犯可能不喜歡與其他人往來，最親近的朋友只有一人，因此也會有沉沒橋接式連結的出現。

接著我們先觀察其他三種社會網路，也就是工廠員工遊戲關係圖、文化夏令

營參加者關係圖和美國南部某鄉村婦女關係網路。這三種網路屬於開放性的朋友網路，網路中的朋友關係不容易受到地形、時間或是其他外在因素影響。一般來說在社會網路中社群的形成會使得群聚度相當高，由圖 22 也可以由強鍵式連結在網路中所佔的高比例得到相同的結果。此外，值得注意的是這三種社會網路和囚犯網路在連結類型的分佈上有些許的不同，例如囚犯網路中的強鍵式連結所佔的比例相對較其他三種網路來的低，這是因為囚犯活動範圍的限制，社群的形成較分散，無法構成大型的社群，使得群聚度相對於其他三種朋友網路來的低。

表 1：四種社會網路的拓樸資訊

網路名稱	節點數	連結數	對應隨機網路的平均網路直徑	分隔度	群聚度
工廠員工遊戲關係網路	14	28	2	3.14	0.83
夏令營參加者關係網路	18	35	3	4.55	0.56
南部婦女活動關係網路	18	50	2	2.68	0.72
囚犯朋友網路	67	182	4	5.8	0.30

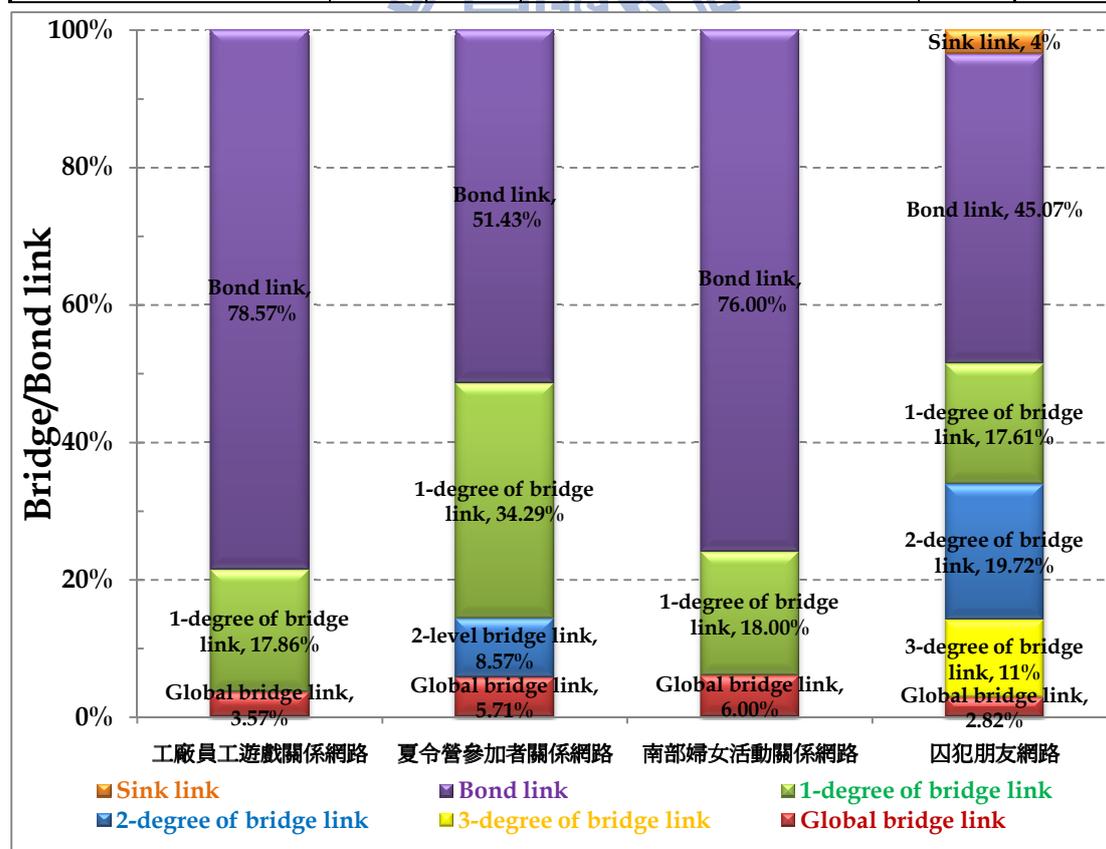


圖 22：四種社會網路連結種類分佈結果統計圖。

4.2.2 生態網路

在生態網路當中，我們選擇了七種不同的食物鏈網路(Williams & Martinez, 2000)，其中有河川或湖泊棲息地，包含了三種食物鏈網路：Skipwith Pond、Little Rock Lake 以及 Bridge Brook Lake；河口棲息地，包含了兩種食物鏈網路：Chesapeake Bay 和 Ythan Estuary；陸生動植物，包含了兩種食物鏈網路：Coachella Valley 和 the island of St Martin(其拓撲性質見表 2)。已經有許多研究指出，生態網路中的弱連結能夠幫助生態平衡遭到破壞或是攻擊時，能夠抑制這些破壞對於網路所造成的損害(McCann et al., 1998; Bascompte, Melian & Sala, 2005; Berlow, 1999)；生態網路中的強連結則容易使得營養瀑布(trophic cascades)所產生的影響變大；然而以往在研究生態網路的動態時，都是利用獵食者-獵物之間的消費關係來定義或量化連結強度，我們的研究則是以拓撲的角度來分析生態網路的網路架構和網路動態的關係。

Dunne 等人指出在大部分的食物鏈網路有類似於小世界網路的低分隔度，但是卻具有異於小世界網路的低群聚度(Dunne, Williams & Martinez, 2002)。圖 23 是七種食物鏈網路經過判斷演算法之後的結果統計圖。可以很清楚的發現，強鍵式連結所佔有的比例都偏低，這將會使得食物鏈網路的群聚性降低，這樣的結果和 Dunne 等人的實驗結果不謀而合。

接著我們考慮食物鏈網路的拓撲架構，藉由 Bascompte(2005)等人的研究中提出在食物鏈網路中的結構要素(Tritrophic food chain, TFC)(Bascompte et al., 2005)，如圖 24。生態網路中的雜食性獵食者的存在會使得拓撲架構產生封閉三角形，會讓群聚度上升。在 Dunne 等人的實驗中證實 Coachella Valley 和 Skipwith Pond 這兩種食物鏈網路的性質被網路中的雜食性物種所影響，因此具有相似性。這些雜食性物種的連結數較多，能夠造成較高的群聚度(Dunne, 2002)。

此外，Chesapeake bay 這個食物鏈網路的分隔度相較於其他六種食物鏈網路的分隔度高，在經過演算法之後，我們可以發現這個食物鏈網路沒有全域橋接式連結，這也說明了全域橋接式連結的存在對於整體複雜網路的分隔度有顯著的影響。

藉由實驗結果顯示，在七種食物鏈網路當中「一階橋接式連結」所佔有的比例是最多的，我們可以藉由圖 25 來分析這樣的結果。食物鏈網路具有明顯的階層性，也就是所謂的營養層級(Trophic levels)，上層物種以下層物種為食物，最上層物種稱之為高級消費者，最下層物種稱之為生產者。圖 25 是 Bridge Brook Lake 的食物鏈網路經過演算法判斷之後的結果，我們可以發現雖然食物鏈網路具有明顯的階層性，能夠使得網路的群聚度降低，連結兩端節點的共同朋友數較少；但是在次級消費者的物種之中的同一階層的物種互相獵食或是跨階層獵食的狀況較為明顯，會使得連結出現的較緊密，造成連結兩端的共同朋友數較多。因此在這兩種情況的交互作用之下，連結的類型不會是強鍵式連結也不會是全域橋接式連結，而是介於其中的 n 階橋接式連結。由圖 25 可以發現大部分的一階橋接式連結(藍色線段)都出現在介於最高級消費者與生產者之間的次級消費者之間。

表 2：七種食物鏈網路的拓樸資訊

網路名稱	節點數	連結數	對應隨機網路的平均網路直徑	分隔度	群聚度
Bridge brook lake	25	104	2	3.92	0.31
Chesapeake bay	31	67	4	5.32	0.21
Coachella valley	29	243	2	2.31	0.69
Little rock lake	92	984	3	3.30	0.47
Skipwith pond	25	189	2	2.08	0.61
St. Martin	42	205	3	3.25	0.32
Ythan estuary	83	391	3	3.32	0.24

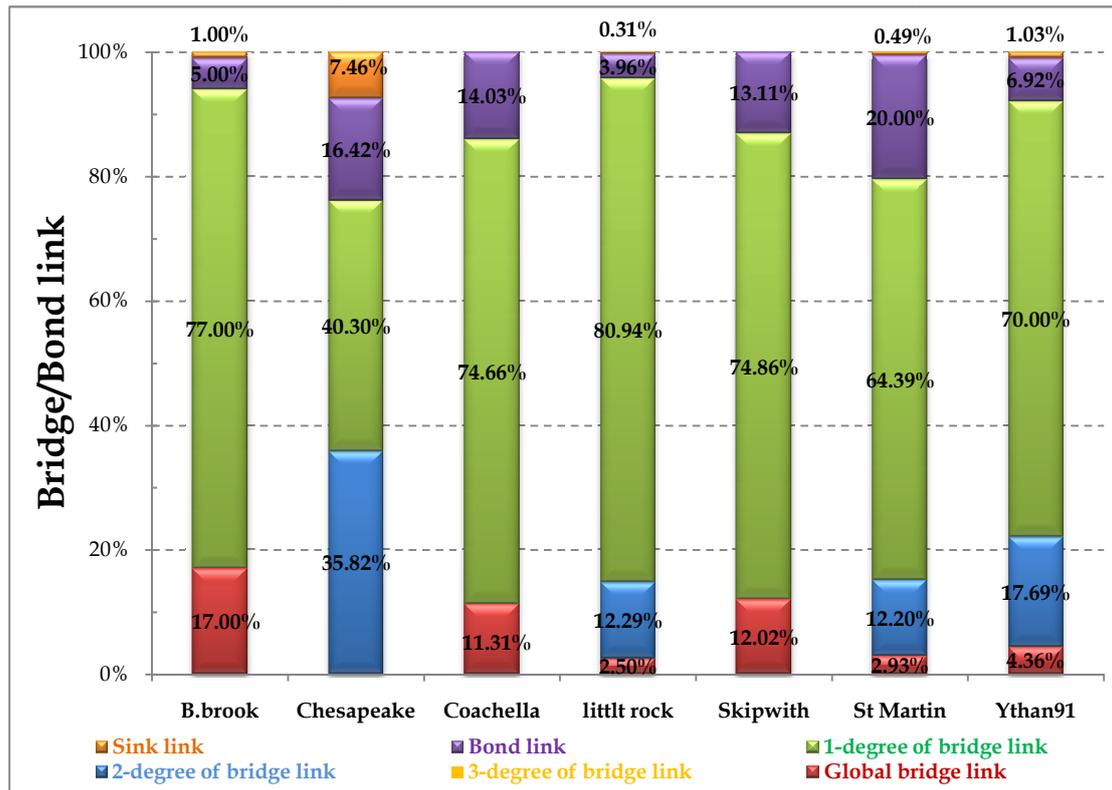


圖 23：七種食物鏈網路連結種類分佈結果統計圖。

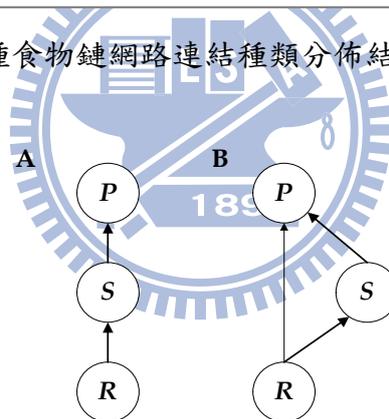


圖 24：Trophic food chain(TFC)。P 表示獵食者(Predator)，S 表示消費者(Consumer)，R 表示生產者(Resource)。A)表示一般 TFC。B)表示 TFC 中擁有雜食性獵食者。

此外，若是網路中的某物種只捕食同一階層的物種或是被同一階層的物種所捕食，這樣的捕食關係較不容易形成共同朋友，那麼這些捕食關係很可能就被判斷成全域橋接式連結。舉例來說，圖 25 中 Bridge Brook Lake 食物鏈網路的最高級消費者與生產者物種的消費關係非常單純。最高級的消費者只捕食次高級的消費者，大部分的生產者只被初級消費者捕食，因此全域橋接式連結大部分都出現在最高級消費者和生產者的消費關係之中。

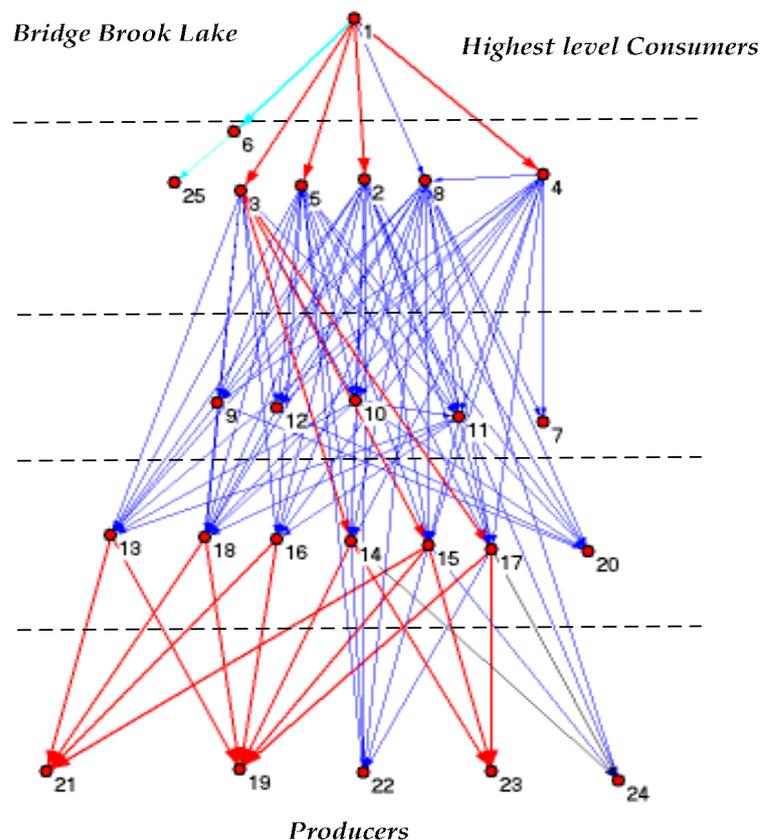


圖 25： Bridge Brook Lake 的食物鏈網路示意圖。紅色連結代表全域橋接式連結。藍色連結代表一階橋接式連結。黑色連結代表強鍵式連結。淡藍色連結代表沉沒式連結。

4.2.3 電子電路網路

在電子電路網路當中，我們測試了 ISCAS'89 基準時序電路(ISCAS89 benchmark set of sequential logic electronic circuits)，包括了 s208、s420 和 s838 三種小數乘法器(拓樸資訊見表 3)(Cancho, Janssen, & Sole, 2001)。電子電路網路中，節點代表邏輯閘與 D 型正反器，連結代表邏輯閘或正反器間的實體電路。基於成本的考量，工程師會盡量簡化電路已達到最有效率的設計。因此在設計電路時，若是能夠維持同樣的功能，那麼以最少的電子元件(電阻、電容或是二極管)為設計的優先考量。因此，工程師會盡量減少冗長以及多餘的電路，以符合電路簡化的設計原則；然而，這樣的設計原則會使得複雜網路呈現較高分隔度與較低的群聚度(Cancho et al., 2001)。

工程師在設計電路時，會存在由較小的「邏輯單元」組成「功能單元」執行

單一任務，這些邏輯單元會緊密地連結在一起，因此功能單元內的連結會是強鍵式連結。「功能單元」會建構成「模組」執行較複雜的工作，而功能單元之間的連結就會是 n 階橋接式連結。換句話說，電子電路網路具有明顯的階層架構 (Berlow, 1999)。圖 26 是三種複雜網路經過演算法計算後的結果，我們可以發現在電子電路網路當中，沒有一階橋階式連結的存在。也就是說，除非某連結在第一階層就被判斷成強鍵式連結之外，否則此連結兩端節點的共同朋友比例一定會通過第一階層的內部門檻。這樣的結果顯示在同一個功能單元內的邏輯單元(節點)之間連接的相當緊密，連結兩端節點在延伸一步之後的共同朋友比例很高，屬於強鍵式連結；而不同功能單元之間的連結兩端節點的共同朋友比例就較低，屬於 N 階橋接式連結。

表 3：三種電子電路網路的拓樸資訊

網路名稱	節點數	連結數	對應隨機網路的平均網路直徑	分隔度	群聚度
S208	122	198	7	8.72	0.059
S420	252	399	8	10.12	0.056
S838	512	819	9	12.37	0.054

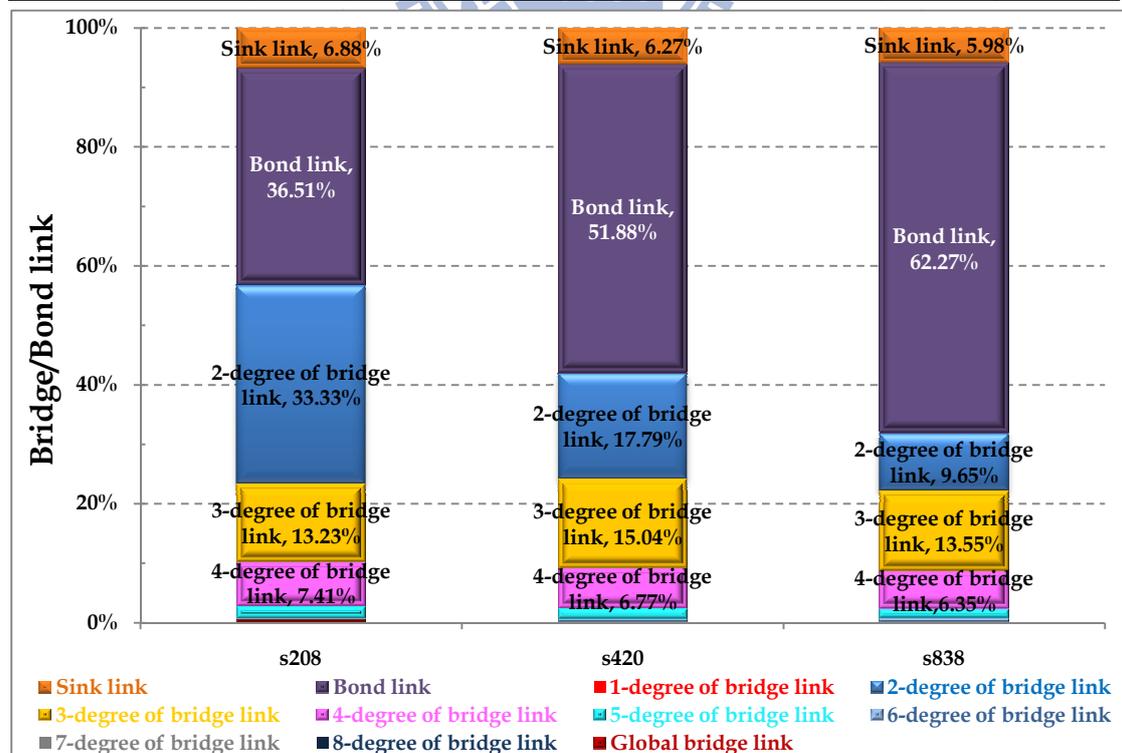


圖 26：三種電子電路網路中連結種類分佈結果統計圖。

在電子電路的設計當中，必須考慮電路板上電子元件之間的相對位置關係，實體電路越長，那麼訊號越容易受到雜訊干擾，因此不容易出現連接兩個較遠距離的電子元件的實體電路，這個情況使得電子電路的拓樸上幾乎沒有全域橋接式連結，這也呼應了電子電路網路的分隔度較高的現象。

圖 27 是電子電路 s208 經過演算法判斷之後所產生的網路拓樸圖，我們可以發現黑色連結會形成完全圖(三角形)，紫色連結會出現在黑色連結附近，並且連結黑色連結所形成的完全圖。藍色連結則會將黑色與紫色連結所形成的區塊連結起來。最後紅色連結則會跨接所有顏色的連結所形成的區塊。以拓樸的角度來分析，黑色連結和紫色連結形成的電路區塊可能是電路中的「功能單元」；藍色則是功能單元之間的連結，能夠將功能單元建構成「模組」；紅色連結則是將模組與零星的電子元件連接成完整的電路；淡藍色連結則是電路的輸入端或是輸出端的電路。許多 ISCAS'89 時序基準電路並沒有提供電路架構示意圖(Brglez, Bryan & Kozminski, 1989)，我們只能夠取得電路的拓樸圖以及電路中電子元件的種類與個數。藉由我們的演算法的幫助，我們能夠提供以拓樸架構為基礎的推測，描繪出電子電路的階層架構。

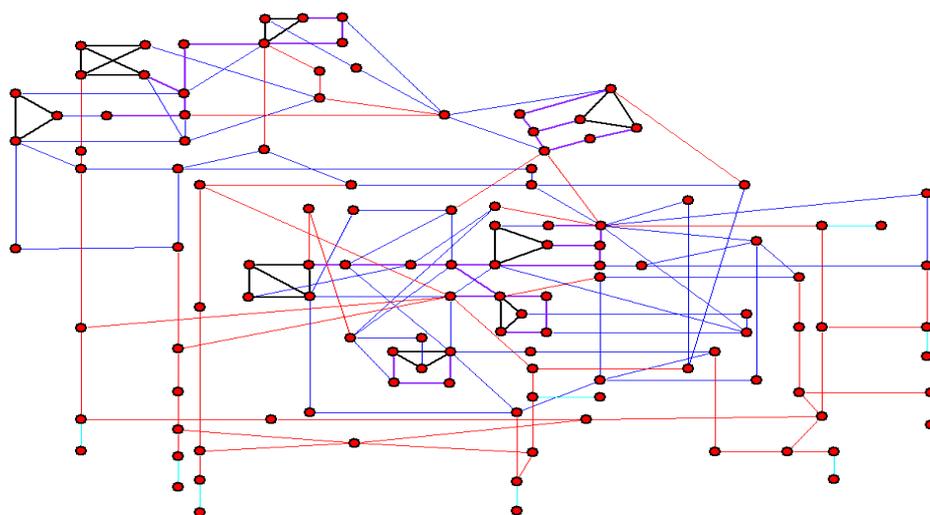


圖 27：s208 的電子電路網路拓樸架構示意圖。黑色連結代表第一階層就被判斷為強鍵式連結的連結。紫色連結代表二階橋接式連結在第三階層轉為強鍵式連結。藍色連結代表二階橋接式連結。紅色連結代表三階或三階以上橋接式連結。淡藍色連結代表沉沒橋接式連結。

第五章 演算法穩固分析與應用

藉由前面的章節我們已經可以瞭解橋接式與強鍵式連結和現有拓樸性質(分隔度和群聚度)的關係，除了能夠讓我們得知橋接式與強鍵式連結的基本特性之外，也對於本研究所提出的演算法做了初步的驗證。此外，在真實世界的複雜網路中各種不同類型的連結會有不同分佈比例，不同類型的連結分佈上的差異也對應到各種複雜網路所展現的動態行為上的不同。舉例來說，電子電路網路中功能單元內的電子元件間的連結較為緊密且功能單元之間的連結較為稀疏，使得電子電路網路中不存在一階橋接式連結；囚犯的朋友網路中，朋友關係取決於地理位置，同牢房的囚犯互相成為好朋友的可能性很高，使得強鍵式連結為此網路中分佈比例最高類型的連結。



當在研究不同領域的複雜網路時，所蒐集到的資料也許不是非常完整，很可能會有錯誤或是失真的狀況。舉例來說，社會網路的問卷調查結果很可能就會因為受訪者的填寫錯誤或是資料遺失而造成資料上瑕疵；生態網路在紀錄物種間捕食狀況時，假設物種甲有物種乙、物種丙和物種丁此三類物種當作獵物，但是物種甲主要捕食物種乙和物種丙，而很少捕食物種丁。在這樣的情況之下，物種甲和物種丁的消費關係很容易就被忽略，造成資料蒐集上的不完全。因此這個章節對於本研究提出的演算法的穩固性做了一連串的實驗，觀察此演算法當面臨不完整的資料時，是否依舊能夠判斷出每條連結正確的性質。

最後，「複雜網路的分群」以及「尋找網路中的社群結構(community structure)」是複雜網路研究範疇中重要的兩個應用。章節 5.2.1 利用演算法所得的連結類型分佈比例當作分群指標，將拓樸結構相似的網路歸類在同一群。章節 5.2.2 利用本研究所提出的演算法所判斷出的連結類型，依據橋接式與強鍵式連結的特性能夠讓我們找尋網路中可能存在的社群。

5.1 穩固分析

本節利用三種類型的網路：社會網路(空手道俱樂部成員關係網路和美國大學美式足球校隊對戰網路)，生態網路(Coachella Valley)和電子電路網路(s420)共四個複雜網路，用來當作演算法穩固分析的對象(拓樸資訊請參考附錄)。針對每一網路，我們隨機移除網路中 $x\%$ 的連結($x=1\%、2\%、3\%、4\%、5\%、10\%、20\%、30\%、40\%、50\%$)，接著針對移除連結後的網路(失真網路)，執行演算法並且得到結果。最後將這些結果與原始資料所產生的結果做比對，觀察移除連結前後的結果差異的變化。

圖 28 是四個網路的比較結果，發現無論是何種失真網路，與原始完整網路所產生的結果的比對正確率和移除連結的比例($x\%$)呈現線性關係。也就是說，當移除的連結數量越多，經過演算法計算後的結果誤差也會跟著提昇。但是即使移除網路中 50% 的連結也不會發生比對正確率突然大幅下降的情況，因此可以看出我們演算法具有一定程度的穩固性。

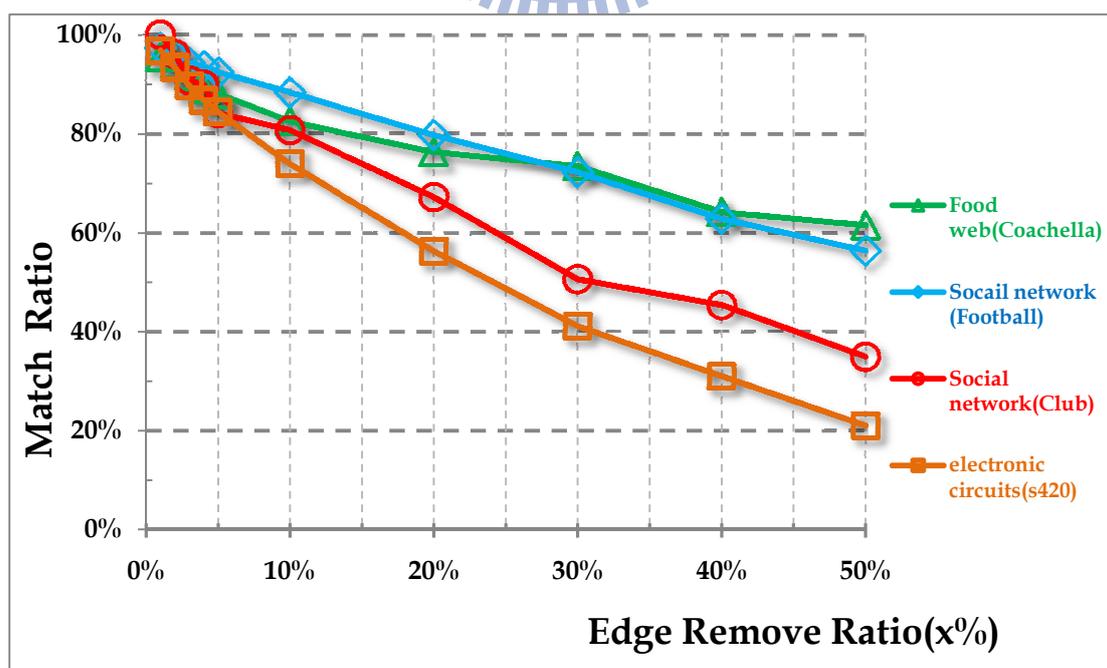


圖 28：演算法判斷失真網路的比對正確率統計圖。

由圖 28 也可以發現不同的複雜網路面對資料不齊全時，比對正確率下降的幅度會有不同。換句話說，不同拓樸結構的複雜網路面對資料失真時，演算法判斷的正確性會受到影響。因此在資料失真的情況下，我們希望能夠瞭解演算法在判斷不同類型的連結時，是否會受到連結類型本身特性的影響。我們將複雜網路中的連結分為三種類型來討論，第一種類型為在第一階層就被判斷為強鍵式連結的連結，特別稱之為「一階強鍵式連結」(bond link of layer 1)，此種連結是網路中最緊密的連結，也是可取代性最高的連結。第二種類型為「全域橋接式連結」(global bridge link)，此種連結是網路中可取代性最低的連結。第三種類型為「其他連結」(else link)，這些連結的可取代性介於全域橋接式連結與一階強鍵式連結之間。

針對「一階強鍵式連結」的基本性質來討論，若是連結被判斷為一階強鍵式連結，表示此連結的共同朋友比例大於外部門檻值，因此當移除連結的比例較低時($x < 10\%$)並不會影響一階強鍵式連結的判斷。當移除連結的數量較多時($x > 30\%$)，整體網路中連結的共同朋友比例也都會下降，也就會直接影響外部門檻值降低，而原本被判斷為一階強鍵式連結的連結還是具有相對較多的共同朋友比例，因此大部分仍舊會被判斷為一階強鍵式連結。也就是說，無論連結移除的比例多寡，並不會嚴重影響演算法對於一階橋階式連結的判斷。圖 29 是在面對移除連結事件的情況下，演算法判斷一階橋接式連結的比對正確率。我們可以發現即使移除了 50% 的連結，比對正確率仍就保持一定的水準，這樣的判斷結果和一階橋接式連結的基本拓樸性質是相對應的。

針對「全域橋接式連結」的基本性質來討論，若是連結被判斷為全域橋接式連結，表示此連結的共同朋友比例相較於網路中其他連結低很多，當移除連結的事件發生時，只會使得網路中連結數更少，而原本就是全域橋接式連結的連結性質不會有所改變。因此，在此在全域橋接式連結沒有被移除的狀況之下，無論連

結移除的比例多寡，演算法在判斷全域橋接式連結時，均不會受到嚴重的影響。圖 30 是在面對移除連結事件的情況之下，演算法判斷全域橋接式連結的比對正確率。我們可以發現其比對正確率下降速度較快，出現這樣的結果並不是因為面對資料失真時，演算法在判斷全域橋接式連結的誤差，而是因為複雜網路中的全域橋接式連結所佔全部連結的比例很低，造成百分比變化程度較大，所以當有些許誤差時，比對正確率就會快速下降。

至於「其它連結」屬於 N 階橋接式連結(n degree of bridge link)，其性質介於一階橋接式連結與全域橋接式連結之間，當面對移除連結事件時，連結類型所佔有的比例會影響比對正確率高低。圖 31 是在面對移除連結事件的情況之下，演算法判斷其他連結的比對正確率。我們可以發現不同網路的比對正確率的下降趨勢有明顯的不同。其中食物鏈網路中的「其他連結」就是一階橋階式連結，佔了網路中全部連結的 74.66%，因此即時移除了 50% 的連結，演算法對於食物鏈網路中其它連結的判斷還是維持一半以上的正確率。反觀電子電路網路中的「其它連結」包含了二階橋接式到六階橋接式連結，而各階橋階式連結分別所佔整體網路中全部連結的比例都很低，所以演算法在判斷其它連結時的比對正確率會快速下降。藉由這個實驗，我們除了能夠說明本研究所提出的演算法具有穩固性之外，我們也能夠瞭解在資料失真的情況下，不同類型的連結對於演算法判斷上會具有什麼樣的影響。此外，除了連結性質本身對於演算法的影響之外，我們可以發現網路中佔有比例越高的連結類型，演算法在判斷時越不容易受到網路失真的影響。

在真實世界當中，真實網路中的連結很可能會因為各種原因而消失，譬如說生態網路中因為環境的變化造成原本有消費關係的兩物種無法維持捕食行為，亦或是電力線網路的攻擊事件很可能造成某些實體電路無法作用。我們可以想像這些事件發生在不同的網路當中會造成不同的影響。然而，在不同的網路中，連結

類型的組成結構也有所不同，舉例來說，社會網路中強鍵式連結較多，其他類型連結較少；生態網路中一階橋接式連結較多，其他類型連結較少。依據本研究提出的演算法的概念，網路中的連結種類越多，表示網路的分隔度越高，當某些連結消失時，會使得剩下的連結較容易發生拓樸性質的改變，使得演算法在判斷上產生不同的結果，造成比對的正確率下降。因此，藉由這個實驗我們認為分隔度越低的複雜網路，在面對網路中連結消失的事件發生時，連結較能夠維持原來的拓樸性質。

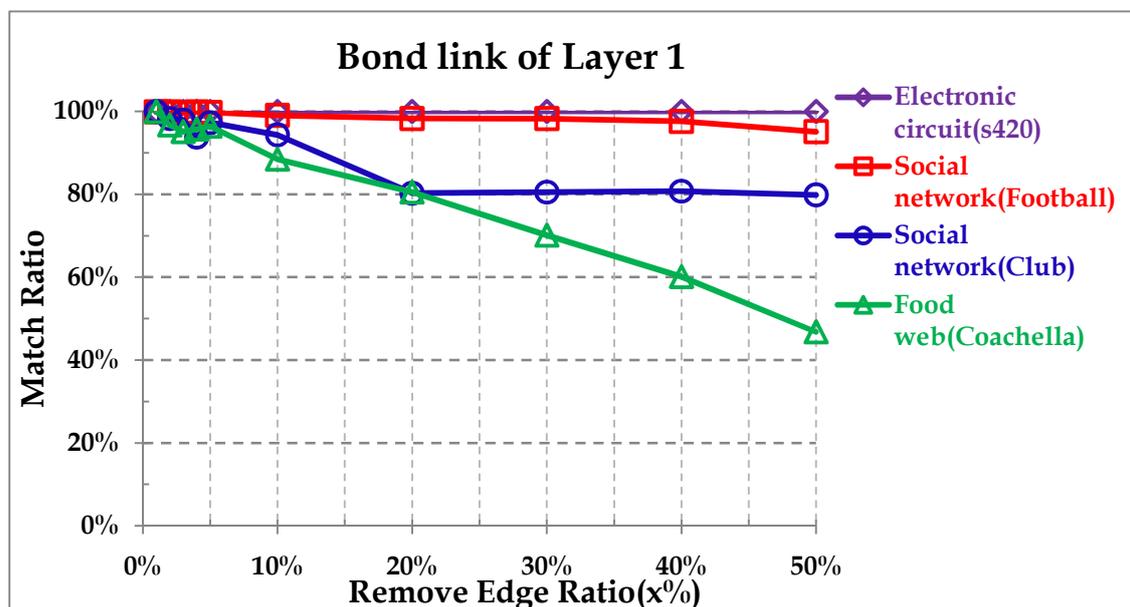


圖 29：移除連結事件下，「一階強鍵式連結」的比對正確率。

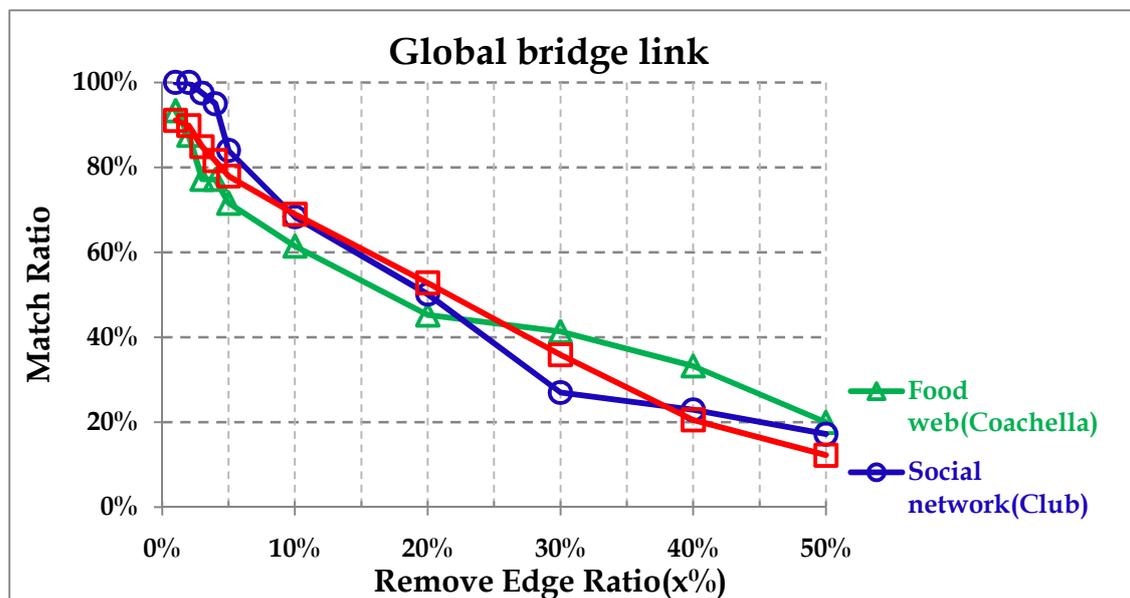


圖 30：移除連結事件下，「全域橋接式連結」的比對正確率

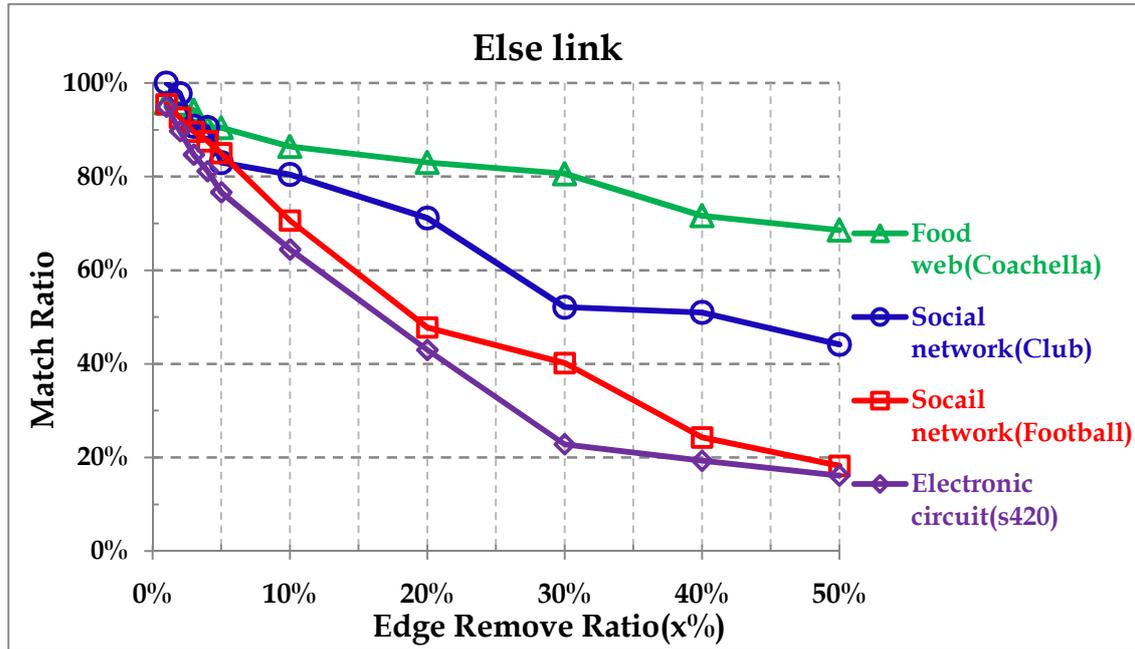


圖 31：移除連結事件下，「其他連結」的比對正確率



5.2 演算法應用

本研究提出的演算法針對複雜網路的偵測結果，我們可以得到網路中具有哪些連結類型以及其分佈的比例，利用這些資訊我們可以在研究複雜網路領域上提供許多的應用。章節 5.2.1 利用連結類型的分佈比例當作分群指標，提出一套複雜網路分群法，能夠將不同性質的網路分群，並且也將真實網路和理論網路做比較，觀察真實網路和何種理論網路在拓撲結構上較為相似，提供往後的研究學者在研究真實網路時，在無法取得真實資料的狀況下，能夠以相似的理論網路當作分析對象或是模擬資料。章節 5.2.2 利用連結類型的性質，本研究依據網路經過演算法偵測後的結果提供了一套簡單的步驟，能夠在找尋網路中的社群結構上的應用提供相當不錯的結果，並且利用了尋找社群結構中的經典例題：「Zachary 的空手道俱樂部成員網路」當作分析對象，來展示本研究和 GN 演算法在此類應用上的差別(Girvan & Newman, 2002)。

5.2.1 複雜網路的分群

Milo 等人在 2004 年利用網路中的三節點與四節點基調(3-node and 4-node motifs)當作分群的指標(Milo et al., 2004)，將具有相同顯著性的基調的網路分為同一群，稱為 Superfamily(參考章節 2.1)。在第四章的真實網路分析結果當中，我們可以發現不同領域的複雜網路會因為拓撲結構上的不同造成連結種類分佈比例也有明顯的差異，因此我們認為連結類型的分佈比例會是複雜網路分群應用上的重要參考資訊之一。本研究利用演算法所得到的連結類型分佈比例中的七項數據當作分群指標，下面是七種分群指標的分群觀點：

1. 一階強鍵式連結之比例：此類型連結屬於網路中最緊密的連結，其分佈的比例與網路群聚度有直接的關係，一階強鍵式所佔比例越高，則網路群聚度就越高。因此，我們將一階強鍵式連結所佔比例當作分群指標之一。

2. 全域橋接式連結之比例：此類型連結屬於網路中可取代性最低的連結，對於網路的資訊傳遞有極大的影響力，並且存在此類型連結的網路都具有相對較低的分隔度。因此，我們將全域橋接式連結所佔比例當作分群指標之一。
3. 沉沒橋接式連結之比例：此類型連結屬於網路中邊緣節點與其他節點之間的連結，若是沉沒橋接式連結所佔比例越高，表示邊緣節點越多，網路拓樸可能會越接近星狀網路。因此，我們將沉沒橋接式連結所佔比例當作分群指標之一。
4. 一階、二階和三階橋接式連結之比例：每一種複雜網路的分隔度均不相同，但是此三種類型的連結所佔比例之總和在大部分的網路中超過 60%。一階橋接式連結的可取代性較高，三階橋接式連結的可取代性較低。因此，我們將這三種類型的連結分別當作三種分群指標。
5. 其他類型連結之比例和：若是網路的分隔度超過三，就會出現四階或四階以上的橋接式連結。一般來說，三階以上的橋接式連結所佔比例都較低，因此我們將四階或四階以上的橋階式連結視為同一類型連結所佔比例，將其比例加總，稱之為其他類型連結之比例和。因此，我們將其他類型連結之比例和當作最後一個分群指標。

接著我們針對第四章中 14 個真實世界網路以上述七個分群指標做分群。如圖 32 到圖 34 所示，同樣領域的複雜網路在連結類型的分佈比例上都呈現高度的相關性，在四個社會網路中，因為群聚現象的關係，一階強鍵式連結比例較其他指標來的高出許多，彼此在分群指標之間的相關係數介於 0.89~0.99 之間；在七個食物鏈網路中，因為網路具有明顯的階層性以及跨階層或同階層消費的現象，兩種性質互相拉扯之下，因此大部分的連結都被判斷為一階橋接式連結，七種網路彼此在分群指標之間的相關係數介於 0.63~0.99 之間；在三種電子電路網路當中，因為電子電路的基本設計原則，使得分隔度較高以及不存在一階橋接式連結，彼

此在分群指標之間的相關係數介於 0.7~0.99 之間。

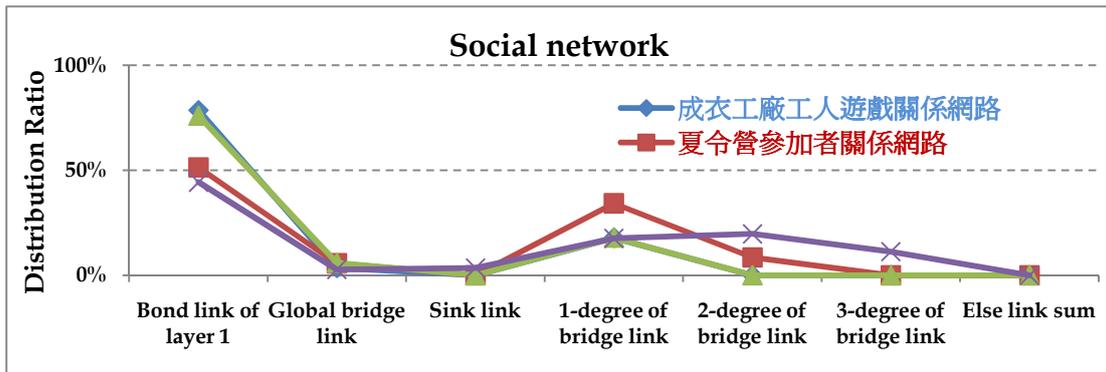


圖 32：四種社會網路在分群指標下的相關性(C=0.9~0.99)。

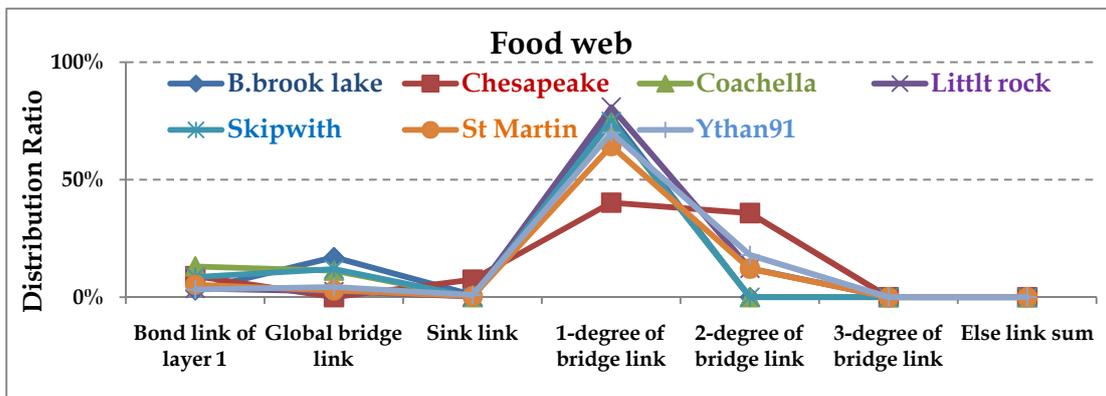


圖 33：七種食物鏈網路在分群指標下的相關性(C=0.62~0.99)。

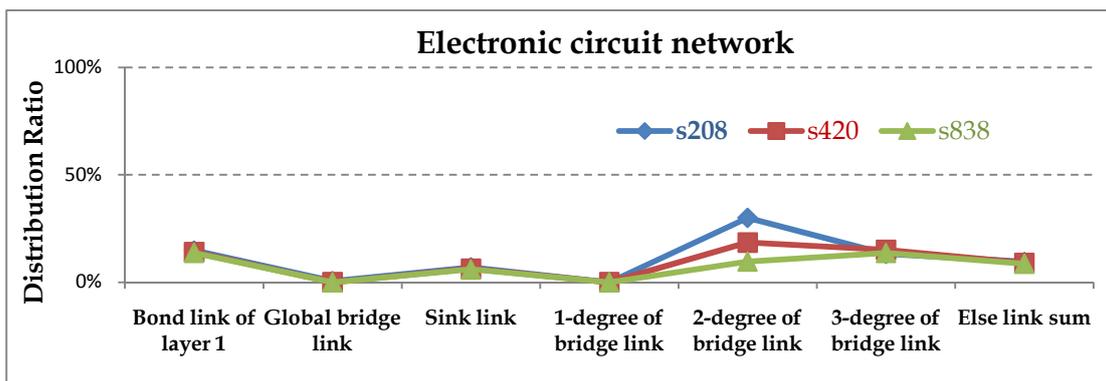


圖 34：三種電子電路網路在分群指標下的相關性(C=0.7~0.92)。

接著，我們兩兩比較 14 種的複雜網路，觀察七種分群指標的相關性，是否會具有顯著的差異，圖 35 是 14 種網路在分群指標的相關係數矩陣視覺圖，由圖

中可以觀察到對角線的黑色區塊(相關係數高)剛好對應到不同領域的複雜網路，而不同領域的複雜網路之間的呈現灰白色區塊。

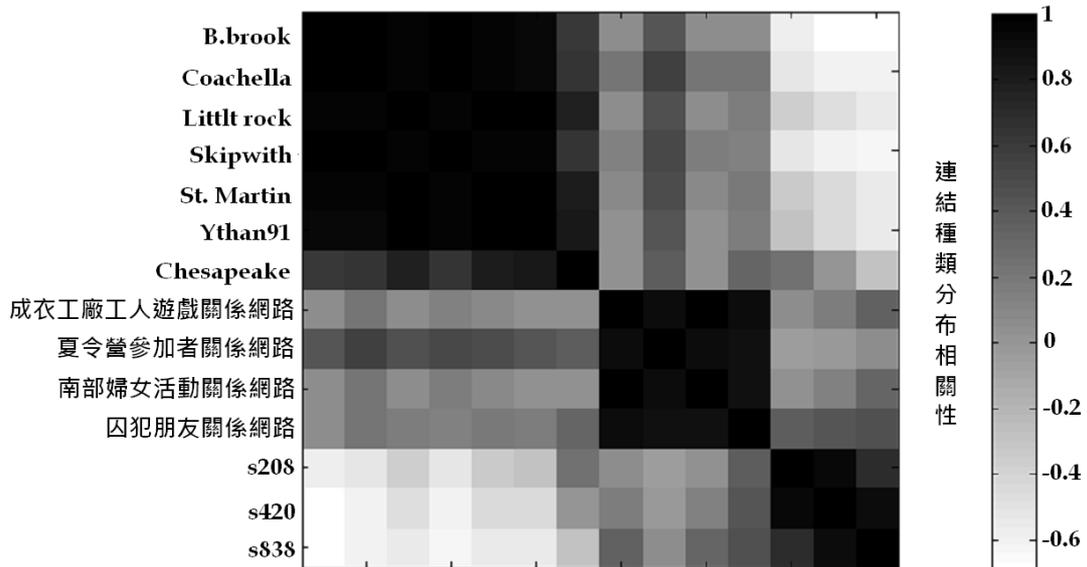


圖 35：14 種網路彼此之間的相關係數矩陣視覺圖。

結果顯示同一領域的複雜網路呈現高度的相關性，而不同領域的複雜網路之間相關性就低的多。這樣的結果清楚顯示出利用連結種類分佈比例當作分群指標具有相當不錯的分群效果。接下來我們利用分群指標將 14 種真實世界網路與 3 種理論網路做比較，理論網路包括小世界網路、無尺度網路和隨機網路(三種網路節點數 $N=100$ ，連結數 $E=400$ ，小世界網路的繞線率 $p=0.04、0.08、0.16、0.32、0.64$)。圖 36 是 14 種真實世界網路和三種理論網路在分群指標上的相關係數矩陣視覺圖，我們可以發現社會網路和小世界網路在繞線率低時($p=0.4\sim 0.16$)在連結類型分佈上具有高度的相關性，並且與無尺度網路和隨機網路的相關性都較低。此外，我們發現以拓樸的角度來說，無尺度網路比隨機網路更接近食物鏈網路的連結類型分佈比例，這樣的結果在以往在探討生態網路的拓樸架構的研究上提供了新的觀點。最後我們發現電子網路不符合任何一種理論網路的連結類型分佈，因為電子網路本身就是一種人工網路，其拓樸性質較為特殊，因此和理

論網路的相關性相當的低。

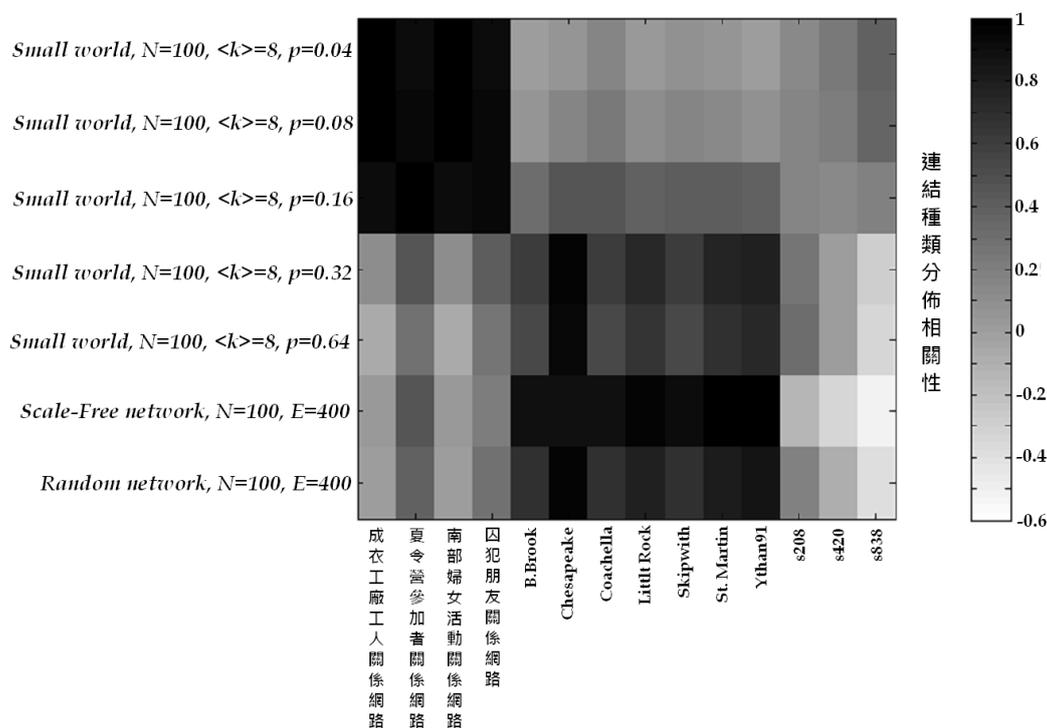


圖 36：14 種複雜網路與理論網路之間的相關係數矩陣視覺圖。

利用連結類型分佈比例當作網路的分群指標能夠提供在做複雜網路的研究時，一個相當有用的工具。舉例來說，當要研究一個未知的複雜網路時，我們能夠先觀察此複雜網路的連結類型分佈比例與哪一種已知複雜網路相似，那麼我們就可以嘗試著用分析已知複雜網路的方式去分析未知的複雜網路，如此一來能夠提高研究效率。

最後，藉由觀察真實網路與理論網路的連結類型分佈比例，我們可以瞭解真實網路與理論網路之間的相關性。往後在各個領域的研究當中，當無法蒐集到真實世界的網路資料時，我們能夠藉由這樣的相關性，利用相似拓模性質的理論網路取代真實網路當作分析對象，能夠降低模擬或是計算上的誤差，得到較為貼近真實世界的分析結果。

5.2.2 找尋網路中的社群結構(Community structure)

隨著對複雜網路的研究發展，很多研究都已經發現許多真實網路都具有一個重要的性質：社群結構(community structure)(Girvan & Newman, 2002)。簡單來說，整體網路架構就是由若干個群體(group)所組成的。在社群結構當中，每個群體內部節點之間的連結相對的緊密，而各群體之間的連結卻較為稀疏。如圖 37 示，這個簡單的網路具有三個群體(虛線包圍的部份)，可以發現群體內的連結非常的緊密，而群體之間的連結就稀疏的多。偵測複雜網路中的社群對於瞭解複雜網路結構與分析是非常重要的。社團結構的分析也在社會學、生物學、物理學以及計算機圖學都具有廣泛的應用。

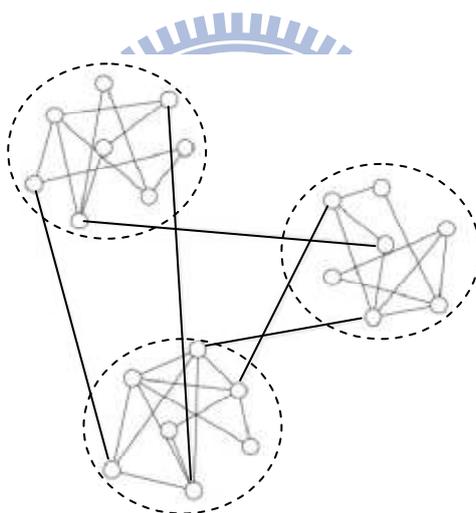


圖 37：具有社群結構的簡單網路。

對於網路社群結構的偵測或判斷，依據針對連結的處理主要可以分成兩大類，分別是分裂法和內聚法。Girvan 所提出的 GN 演算法就是分裂法的代表，GN 演算法的核心概念為不斷移除網路中 Edge-betweenness 值最高的連結，其演算法的基本流程如下：

步驟一：計算網路中所有連結的 Edge-betweenness 值。

步驟二：找到 Edge-betweenness 值最高的連結並且將其從網路中移除。

步驟三：移除最高 Edge-betweenness 值的連結之後，重新計算網路中會受到移

除動作影響的連結的 Edge-betweenness 值。

步驟四：重複步驟二，直到每一個節點就是一個退化的社團為止。

接著我們來看 GN 演算法在社會網路研究中經典問題的應用。此問題是出自二十世紀七零年代初期，Zachary 運用兩年的時間觀察美國一所大學中的空手道俱樂部中內部成員之間的互動關係網路(以下稱作 Zachary 網路)。在他觀察的過程當中，該俱樂部的主管與校長之間為了是否要抬高俱樂部的收費問題產生了爭執。結果，此俱樂部分裂成了兩個分別以主管和校長兩人為核心的小俱樂部(拓樸資訊見附錄 G)。圖 38 節點 1 和節點 33 分別為俱樂部的主管與校長，而三角形和正方形節點也分別代表了俱樂部分裂之後的成員所形成的兩個小型群體。

利用 GN 演算法分析 Zachary 網路，結果如圖 39 所示，若是將 GN 演算法所得到的分裂結果和真實情況做比對，可以發現到其結果與實際分裂的情況幾乎相同。只有節點 3 被判斷錯誤，Girvan 認為節點 3 和兩個群體之間各有五條連結，此節點本身就具有拓樸上的分歧性，因此在演算法判斷上才會出現錯誤。但是 GN 演算法無法用 Edge-betweenness 的結果對應到真實世界去解釋為何會出現錯誤，這樣的情況屬於 GN 演算法在找尋社群結構上的偏見。

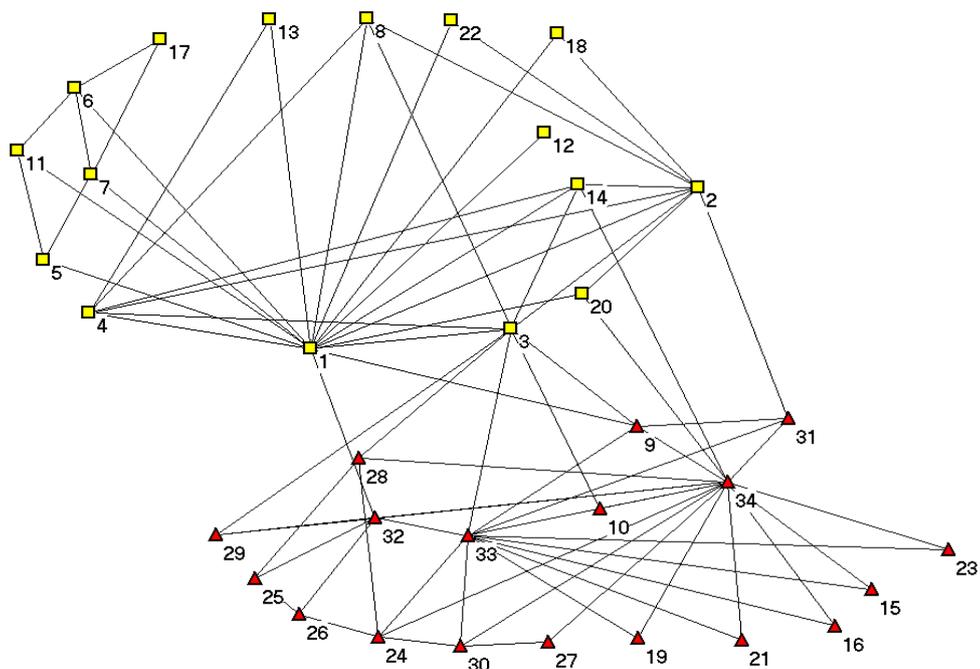


圖 38：Zachary 研究美國某大學空手道俱樂部成員互動網路。

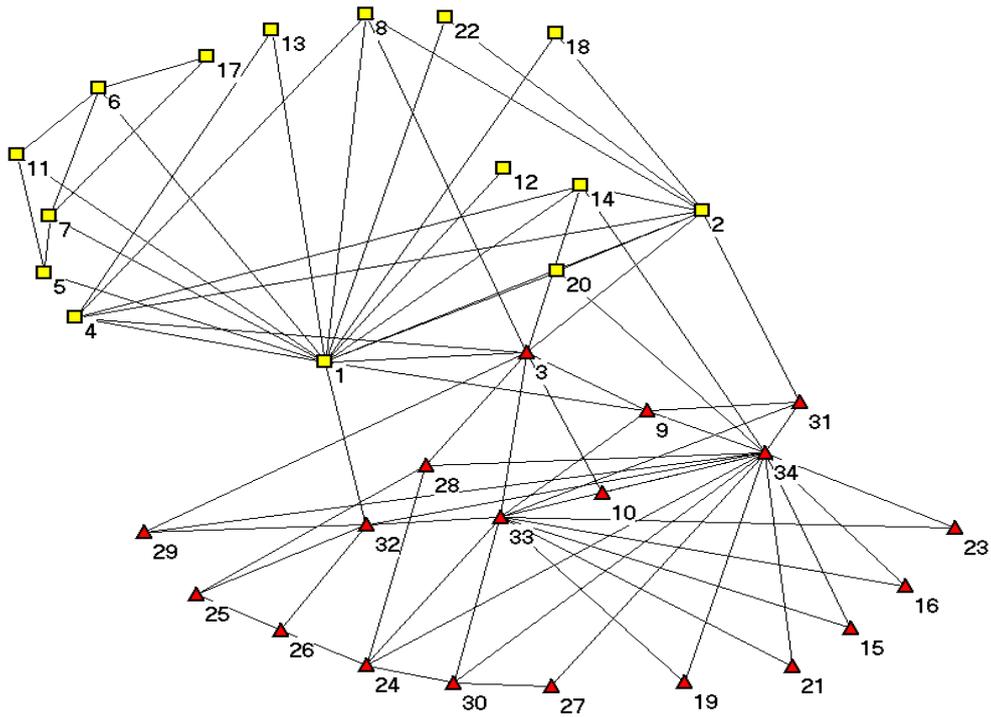


圖 39：GN 演算法分析 Zachary 網路的結果(Girvan & Newman, 2002)。

藉由社群結構的概念，同一個群體內的連結兩端節點的共同朋友比例會相對較高，而群體之間的連結兩端節點的共同朋友比例就會相對較低。舉例來說，社會網路中同一個社團或是群體內的任兩個人若是互相熟識，那麼此兩人在同一個社團或是群體內也會有許多共同認識的朋友；而不同社團或是群體之間的兩人若是互相熟識，那麼他們的共同朋友的數量可能就會少的多。本研究所提出的演算法，能夠依據「共同朋友」的概念將網路中的連結區分成多種不同的類型。因此，我們能夠利用這些不同類型的連結來幫助偵測複雜網路中的社群結構。我們認為強鍵式連結存在於同一個群體內的節點之間；全域橋接式連結則存在於群體之間。

圖 40 是利用本研究提出的演算法分析 Zachary 網路的結果，直接利用不同類型的連結將網路中的社群結構劃分出來。我們提出一個簡單的找尋社群結構的步驟：

初始化：只有節點的網路。

步驟 1：將一階強鍵式連加入網路中。

步驟 2：將 N 階橋接式連結依據共同朋友比例值由高到低加入網路中 (N=1~D, D 是網路的直徑)。

步驟 3：若是 N=D 則演算法結束，否則 N=N+1 回到步驟 2。

藉由上述的步驟能夠依據連結的性質將複雜網路中的節點歸類到正確的群體當中。圖 40 是利用上述步驟所得到的樹狀圖(dendrograms)，可以由圖中發現本研究所提出的演算法計算出來的連結性質來做社群結構的偵測可以得到不錯的成果，與真實情況完全符合(圖 38)。圖 41 是將演算法的結果還原到原始資料的網路拓樸圖，可以清楚觀察到兩個俱樂部之間的連結大部分均為全域橋接式連結與二階橋接式連結，俱樂部內部的連結大部分均為強鍵式連結與一階橋接式連結，經由圖 41 我們更能夠清楚看出本研究提出的演算法所判斷出的連結類型在尋找網路中社群結構的應用中所具有的功能性。另外值得注意的地方是，拓樸性質分歧的節點 3，由圖 41 可以發現它和節點 8 之間具有強鍵式連結，也就是說節點 3 和節點 8 具有較多的共同朋友，因此利用不同的連結類型所具有的性質能夠將節點 3 區分在正確的群體內。

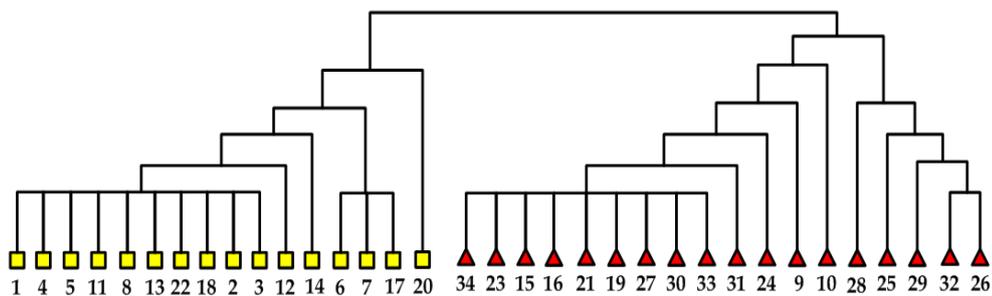


圖 40：經由找尋網路中社群架構步驟所得到的樹狀圖。

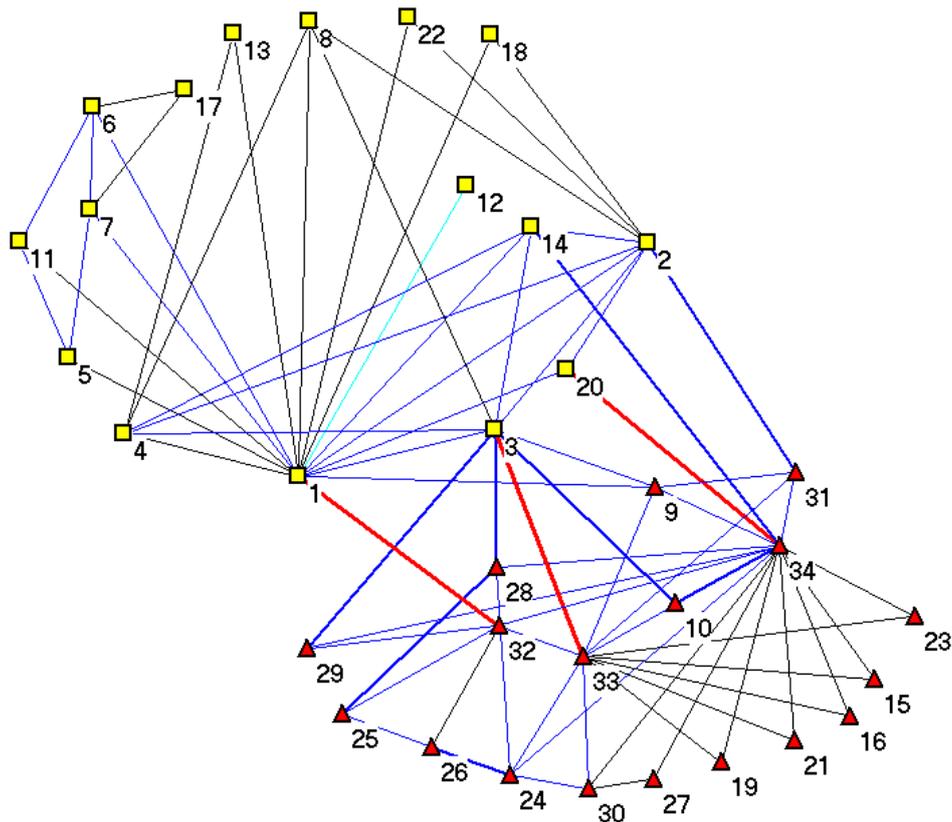


圖 41：本研究所提出的演算法分析 Zachary 網路的結果。紅色連結代表全域橋接式連結。細藍色連結代表一階橋接式連結。粗藍色連結代表二階橋接式連結。黑色連結代表強鍵式連結。淡藍色連結代表沉沒式連結。

以共同朋友的角度來觀察，我們能夠利用演算法的分析結果提供一個簡單的分群步驟達到與現實狀況相符合的結果。此外，針對社會網路的社群結構而言，若是發生與現實世界不符合的情況時，我們還是能夠利用共同朋友的概念解釋分群的結果，以 Zachary 網路為例，假設現實世界當中節點 3 真的屬於紅色三角形的群體當中，也就代表了本研究提出的演算法將節點 3 分到黃色正方形的群體當中是違背現實情況的。但是這樣的結果並不是演算法的錯誤，因為節點 3 和黃色正方形的群體本來就有較多的共同朋友，但是在現實世界當中卻是被紀錄在紅色三角形的群體當中，這種「身在曹營心在漢」情況值得社會學家去探討這樣的人對於社會網路具有什麼樣的影響。

在社群結構的研究上，以 GN 演算法為基礎的偵測方法不斷地被之後的研究

學者所改進，無論是在時間複雜度或是空間複雜度都有明顯的改善。經由分析 Zachary 網路的結果，利用網路中不同類型的連結所具有的性质能夠快速且明顯的勾勒出社群結構的輪廓，說明本研究所提出的演算法可以在偵測社群結構的應用上，提供新的拓樸觀點。



第六章 結論

本研究提出了一個以「共同朋友」為核心概念的演算法，並且利用「階層概念」，能夠將複雜網路中每一條連結依據共同朋友的多寡給予不同的階層性質。連結兩端節點的共同朋友數量越多，表示此兩節點可能在網路中同一個社群之內，當此連結消失時，兩節點的訊息仍然能夠藉由這些共同朋友來傳遞訊息，因此這樣的連結為強鍵式連結；節點兩端節點的共同朋友數量越少，表示此兩節點在網路中屬於不同的社群，當此連結消失時，兩節點的訊息將要花較長的時間才能完成傳遞或是根本無法到達目的地。為了避免以往的研究對於「強弱連結」模糊不清的概念以及非強即弱的二分法，藉由本研究提出的演算法，我們將複雜網路中的連結依據在各階層的「外部門檻值」與「內部門檻值」分類成強鍵式連結(Bond link)、全域橋接式連結(Global bridge link)、N階橋接式連結(N degree of bridge link)和沉沒橋接式連結(Sink link)。瞭解網路中各種類型的連結所存在的比例能夠幫助我們在研究不同的複雜網路時，使我們更進一步瞭解不同性質的網路在拓撲架構上的差異。

實驗結果發現，在第一階層就被判斷成為強鍵式連結的連結數越多，那麼網路的群聚度就會越高，因此一階強鍵式連結在網路中所佔有的比例反應出網路的群聚度。此外，若是網路中具有全域橋接式連結，那麼此網路的分隔度就會相對較低(在本研究分析的社會網路中都具有全域橋接式連結)；若是網路中不具有全域橋接式連結，那麼此網路的分隔度就會相對較高(電子電路網路中均不具有全域橋接式連結)。N階橋接式連結的不可取代性介於強鍵式連結與全域橋接式之間，除了能夠表示連結的區域影響力之外，N值也能夠反應整體網路的分隔度，當N越大，表示網路的分隔度越高。因此，本研究提出的演算法所提供的資訊當中，除了將連結賦予階層特性之外，也能夠反應出網路分隔度與群聚度的基本拓撲性質，不同的連結類型所具有的性質整理如表4。

表 4：不同連結類型性質整理表

類 型 性 質	強鍵式連結	N 階橋接式連結	全域橋接式連結	沉沒橋接式連結
共同朋友數	多(緊密) ←—————→ 少(稀疏) 無			
不可取代性	最低	中等	最高	高
與其他拓撲性質的關係	提高群聚度	/	降低分隔度	無影響
例子： 社會網路	群體之中的連結	不同大小群體之間的連結	群體之間的連結	邊緣個體

從區域觀點到全域觀點瞭解網路中的連結性質能夠針對不同的複雜網路中不同類型的連結提供不同的策略。舉兩個簡單的例子來說，社會網路和交通網路是傳染病模擬的研究對象，以交通網路為例，經過本研究提出的演算法的計算，能夠偵測出網路中哪些連結對於爆發疫情的區域具有擴散傳染的影響力，也就是 N 階橋接式連結所扮演的角色。接著能夠依據演算法的分析結果針對這些 N 階橋接式連結加以管制或是隔離，能夠提高傳染病的傳播門檻。在電力線網路當中，節點代表電塔或是變電箱，連結代表實體電線。以台灣地區為例，新竹科學園區是台灣的經濟重鎮，用電量相當高，當發生跳電或是無法正常供電的情況時，損失是難以估計的。因此，在台灣的電力線網路當中，鞏固新竹科學園區附近的具有區域影響力的連結是個重要的問題。經由本研究提出的演算法，能夠偵測出哪些連結對於新竹科學園區這個區域具有重要的影響力，如此以來，就能夠以較有效率的方式確保科學園區的供電。

真實網路的實驗結果顯示，不同的複雜網路會有不同的連結類型分佈，會反

應出網路不同的動態行為。因此，藉由不同的連結種類與其分佈，本研究在複雜網路中兩個重要的應用面向提出以連結性質為觀點的方法。首先，在複雜網路分群的應用當中，我們利用連結類型的分佈比例當作分群指標，提供一個與網路基調模型(如 Milo et al, 2004)完全不同觀點的分群法，能夠有效的將不同領域的複雜網路區分出來，並且我們也比較了真實網路和理論網路的連結類型分佈比例，我們認為以連結類型分佈的觀點來看，食物鏈網路比起隨機網路來說，更接近無尺度網路的拓樸架構。

接著是在找尋網路中的社群結構的應用當中，我們利用演算法所偵測到的不同連結類型所具有的性質，提出一個簡單的分群步驟，能夠找尋出網路中的社群結構。藉由 Zachary 的空手道俱樂部成員網路當作研究對象，我們能夠將這個網路中最受爭議的節點歸類在正確的社群當中。此外，利用共同朋友的概念，在社會網路當中尋找社群結構時，能夠給予實際的分群意義，而不只是拓樸上特性。以往在「複雜網路分群」和「找尋複雜網路中的社群結構」這兩個研究範疇當中，各自發展出了許多不同的演算法來解決各自的問題。然而，利用本研究提出的演算法所到的拓樸資訊，能夠針對這兩個研究範疇給予一致的解決觀點，並且能夠觀察到前人研究未觀察到的要點，提供往後在研究這些應用問題時能有更多的參考資訊。

連結的拓樸性質階層化的概念，在複雜網路研究領域當中提供了一個全新的觀點，利用不同連結類型的性質與分佈比例除了能夠描述網路的拓樸特徵或是動態行為之外，更能夠在許多複雜網路的應用方面提供更直接的解決方法，這些新的資訊能夠幫助研究學者在研究複雜網路的設計原則或是演化原則上有更進一步的瞭解。

第七章 未來展望

找出網路在複雜架構下隱藏的規則一直都是複雜網路研究當中最主要的目標之一，雖然我們可以用高深的數學公式和計算能力強大的電腦歸納或模擬出複雜網路所展現的行為以及特徵，但是仍有許多未知的特性等著研究學者去探索。本研究利用網路中連結的拓撲特性和架構來描述網路的特性，提供複雜網路的研究學者很多新的資訊，但是本研究仍有不足之處，主要可以分為三個部份來討論，提供給往後的研究學者參考。

第一部份，本研究使用的網路資料均屬於中小型的複雜網路，希望以區域觀點預測全域現象，當網路資料過於龐大時，演算法可能花費的時間將會依據網路中連結數呈現指數成長，若是能夠有效降低演算法的複雜度，那麼就可以針對許多大型網路做分析，並且拉近實驗結果與真實世界的誤差。因此，改善演算法的複雜度將是最主要的議題。

第二部份，考慮複雜網路中連結的方向性與權重。許多複雜網路中個體之間的關係都具有方向性，例如生態網路、基因調控網路和蛋白質網路...等等，若是將本研究提出的「共同朋友」的概念加上連結的方向性，重新定義連結兩端節點的共同朋友所代表的意義，能夠將連結類型依據連結方向性的不同細分成更多不同類型的連結，如下圖 42 所示，圖 42 a 是不考慮方向性時的狀況，對於連結 e 的兩端節點 A 和 B 來說，節點 C 為節點 A 和節點 B 的共同朋友。圖 42 b 是考慮方向性之後的三種情況，此時依據連結的方向性，必須重新考慮節點 C 和節

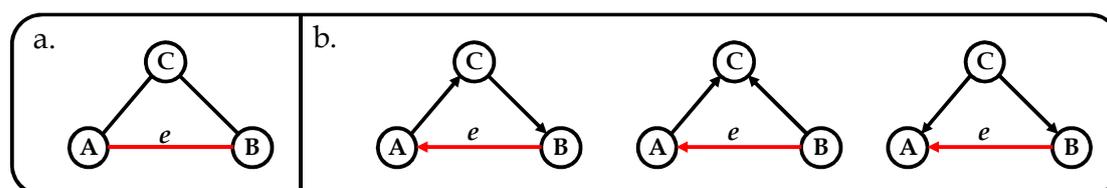


圖 42：共同朋友與連結方向性。a.不考慮方向性。b.考慮方向性。

點 A 與節點 B 之間的關係。此外，複雜網路中個體之間的關係通常不是等價的，舉例來說，在社會網路當中，兩人之間彼此認識，那麼在網路拓撲中就有連結代表兩人的認識關係，但是社會網路中的這些認識關係一定是不相同的，因此很多研究賦予連結權重值來區分個體間關係的差別。利用本研究提出的演算法和權重值的概念加以整合，能夠更精確的判斷出共同朋友在網路中所具有的性質。如圖 43 所示，圖 43 a 表示不考慮權重的情況，換句話說，所有連結的權重均為 1。而圖 43 b 表示考慮權重值的情況，連結 AC 與連結 BC 的權重值均為 0.8。雖然圖 43 a 和圖 43 b 中的節點 C 都代表了節點 A 和節點 B 的共同朋友，但是經由權重值的差異我們可以推斷圖 43 a 中的節點 C 的強度和圖 43 b 中的節點 C 的性質是不相同的。多了方向性與權重值的資訊，能夠使得我們提出的演算法所判斷出的連結類型不再只侷限於拓撲上的性質，而是能夠對應到真實世界所具有的功能性，讓我們能夠更深入的分析複雜網路的行為與現象。

第三部份，在演算法的應用方面還有許多改進空間和延伸研究。舉例來說，複雜網路分群方面，在演算法的結果當中可以找出更多的資訊當作分群指標，以及針對分群指標的正規化計算都能夠幫助分群的準確性以及合理性。

瞭解網路中每一條連結從區域到全域的特性能夠提供各個複雜系統領域的研究學者更多參考資訊，藉由演算法所得的連結類型針對不同的複雜網路具有不同的背景意義，本研究只針對了社會網路、生態網路和電子電路網路當作分析對象，如何將其他更多的複雜網路中具有階層性質的連結對應到真實世界的背景意義，則是要等待未來的研究學者來探索與發現。

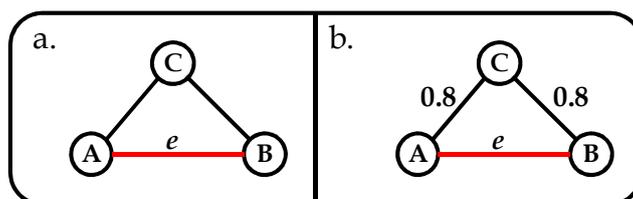


圖 43：共同朋友與連結權重值。a.不考慮權重值。b.考慮權重值。

附錄.A

1. 連結權重值的定義

網路中某一連結 $edge(a, b)$ ，其權重值(weight)可以定義成

$$weight(a, b) = \sum_i \frac{1}{length(path(a, b)_i)}$$

其中 $path(a, b)_i$ 是節點 a 到節點 b 的第 i 條可能的路徑。

$path(a, b)_i \neq edge(a, b)$ 且 $length(path(a, b)_i) \leq Network_average_dimeter$

$$Network_average_dimeter = \frac{\sum_{a,b,a \neq b} ShortestPath(a,b)}{C_2^n}$$

2. 強弱連結的定義

產生隨機網路來計算這些隨機網路之連結的平均權重：

$$(RandNetwork_average_weight)_i = \frac{\sum_{a,b,a \neq b} weight(a,b)}{C_2^n}$$

產生 100 組隨機網路計算出門檻值： $threshold = \frac{\sum_i (RandNetwork_average_weight)_i}{100}$

經過判斷 weight 是否大於或小於此門檻值來判斷其強弱，也就是說：

$weight(a, b) > threshold$ 則 $edge(a, b)$ 屬於強連結

$weight(a, b) < threshold$ 則 $edge(a, b)$ 屬於弱連結

附錄.B

Onnela 等人為了要觀察網路拓樸與連結強度之間的關係與全域資料散播 (Onnela et al., 2007)，他們設計了以下實驗步驟：

1. 在 time 0 的時候我們隨機挑選幾個受到新資訊感染的個體作為初始化。
2. 然後在每一個 time step，受到感染的個體 v_i 就會將資訊傳遞到自己的鄰居(直接連結的節點)，而傳遞的有效機率為 $P_{ij} = xw_{ij}$ (其中 x 為傳播速率)。
3. 在同樣的網路中，將每一個連結的權重用平均權重做代換，因此 P 值都會相同，然後進行同樣的實驗做比較。

由下圖 44.A 可以看出實驗結果，Control 曲線代表了連結權重為平均權重所產生出來的結果，而 Real 曲線則是代表真實世界實際的結果。我們可以發現 Control 的傳播速度比 Real 的傳播速度快很多，這個結果和我們預期的一樣，因為在真實世界當中，複雜網路中的社群會使得資訊傳播產生延遲的現象。

圖 44.B 可以解釋資訊傳播延遲的現象。我們可以看到當資訊剛開始傳播的時候($t < 10$)，資訊會在社群內快速傳播(連結強度較強)，而當社群內的節點都被傳遞到此資訊之後，就會將此資訊傳遞到其他的社群之中，而社群之間的資訊傳遞速度就會相對較慢(連結強度較弱)，所以會有接近水平曲線的傳播速度($10 < t < 20$)。接著傳遞到其他社群之後，這個資訊又會在此社群之間內開始傳播，因此傳播速率又提升上來($t > 30$)。然而在 Control 的情況當中，因為每一條連結的權重均相同，因此社群之間的連結傳遞機率也是和社群的傳遞機率一樣，因此傳遞速度的結果相較於 Real 的情況會快的多。

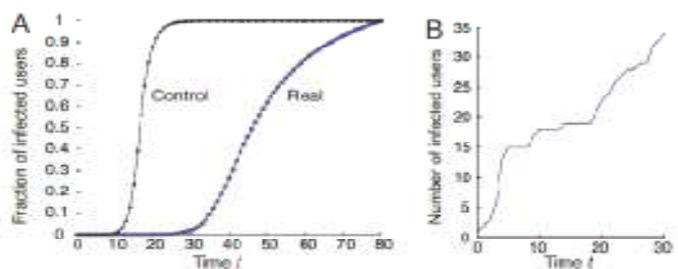


圖 44：傳播速率圖(Onnela et al., 2007)。

附錄.C

Fedderke 等人(1999)主張將社會資本橫切，區分成兩種功能：「透明度」與「理性化」(Halpern, 2008)。透明度指的是一個社群的社會資本如何使資訊流通更加便利。因此，一個擁有較多橋接性社會資本的社會，將被視為擁有更多的透明度，而將資訊更快速地在社群與團體間流通。理性化指的是社會資本從重視實質內容的規則與規範轉變成重視程序的規則與規範之程度。因此，某個社群的社會行為若是被嚴格的傳統規定所管轄，另外一個社群的社會規則是由情境與結果而定，那麼前者的理性化程度較低。圖 45 是強鍵式(黏著性)與橋接式(架接性)社會資本以不同的混和方式創造出不同的社會形態。

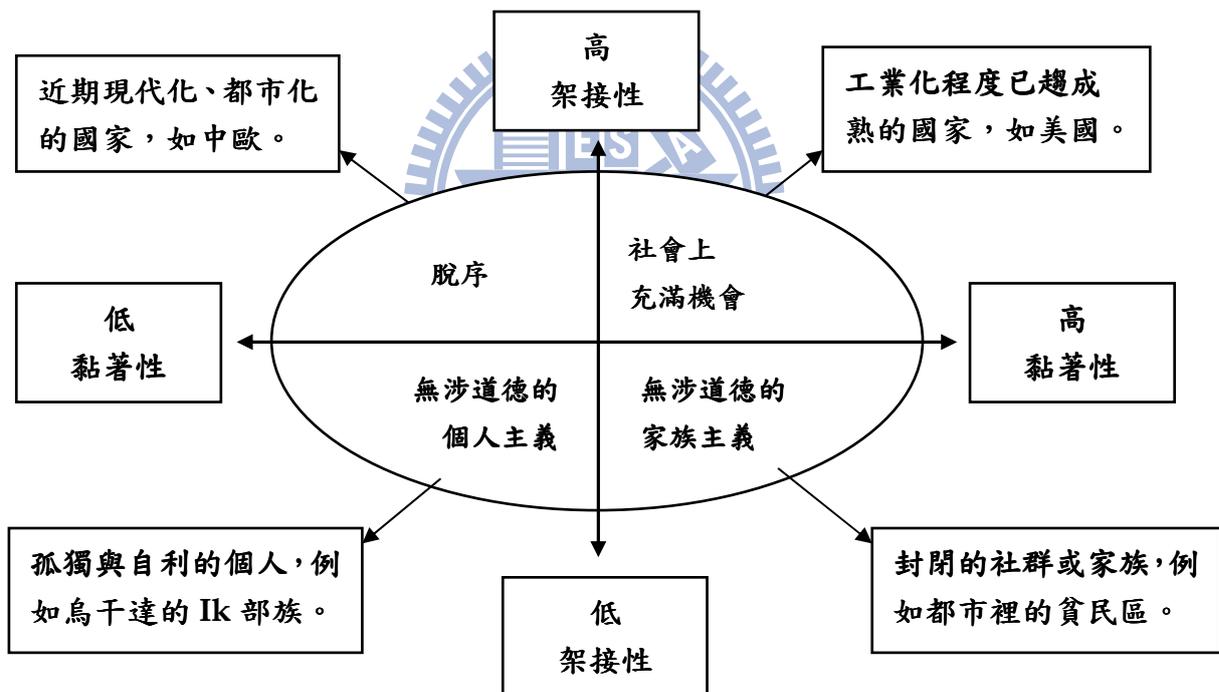


圖 45：不同性質的社會資本混合下所產生不同型態的社會(Halpern, 2008)。

附錄.D

連結兩端節點共同朋友比例

連結 $e = (i, j)$ 的共同朋友比例 O_{ij} 為：

$$O_{ij} = n_{ij} / ((k_i - 1) + (k_j - 1) - n_{ij})$$

其中 k_i 為節點 i 的分支度； k_j 為節點 j 的分支度。

n_{ij} 為節點 i 與節點 j 的共同朋友個數。

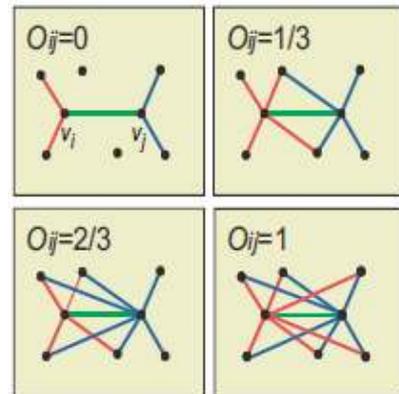
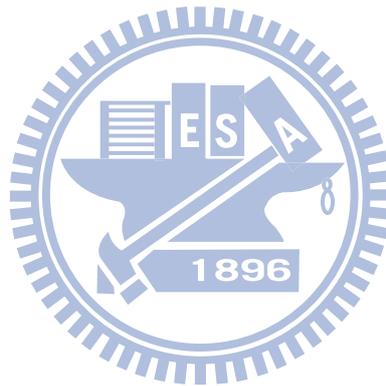


圖 46：共同朋友比例。



附錄.E

WS 小世界網路

表 5 是實驗結果統計圖，利用 Watts 和 Strogatz 所提出的方式產生小世界網路，節點數為 100，固定平均分支度 $\langle k \rangle$ ，觀察繞線機率 p 的變化與橋接式連結和強鍵式連結的分佈之間的關係 ($p=1、2、4、8、16、32、64$)。

表 5：不同繞線機率下的 WS 小世界網路

Category	Edges	Rewiring(p)	$C(p)/C(p=0)$	$S(p)/S(p=0)$	Bridge link	Bond link	Sink link
Small-World #1 Degree $\langle k \rangle = 2$ Nodes=100	100	1%	0.0%	64.7%	21.0%	41.0%	38.0%
		2%	0.0%	100%	28.0%	37.0%	35.0%
		4%	0.0%	31.3%	4.0%	39.0%	57.0%
		8%	0.0%	44.4%	6.0%	41.0%	53.0%
		16%	0.0%	40.47%	59.0%	8.0%	33.0%
		32%	0.0%	24.32%	43.0%	11.0%	46.0%
		64%	0.0%	28.43%	40.0%	18.0%	42.0%
Small-World #2 Degree $\langle k \rangle = 4$ Nodes=100	200	1%	93.87%	58.8%	3.5%	96.5%	0.0%
		2%	94.07%	47.36%	3.5%	96.5%	0.0%
		4%	79.67%	36.24%	11.5%	88.5%	0.0%
		8%	64.47%	28.84%	24.5%	75.5%	0.0%
		16%	46.18%	25.0%	40.0%	60.0%	0.0%
		32%	20.13%	22.38%	67.5%	31.5%	0.0%
		64%	11.46%	21.78%	82.0%	16.0%	0.0%
Small-World #3 Degree $\langle k \rangle = 8$ Nodes=100	400	1%	90.56%	50.85%	3.75%	90.56%	0.0%
		2%	94.05%	52.31%	2.75%	94.05%	0.0%
		4%	88.41%	42.85%	4.75%	88.41%	0.0%
		8%	70.44%	34.62%	15.0%	70.44%	0.0%
		16%	46.67%	30.46%	37.5%	46.67%	0.0%
		32%	26.19%	29.31%	64.5%	29.16%	0.0%
		64%	14.22%	27.23%	88.5%	14.22%	0.0%
Small-World #4 Degree $\langle k \rangle = 16$ Nodes=100	800	1%	97.42%	70.86%	0.75%	99.25%	0.0%
		2%	90.5%	54.71	3.38%	96.62%	0.0%
		4%	84.03%	50.43	5.88%	94.12%	0.0%
		8%	65.19%	42.86	19.5%	80.5%	0.0%
		16%	53.29%	42.86	33.88%	66.12%	0.0%
		32%	35.83%	42.86	70.25%	29.75%	0.0%
		64%	25.49%	42.71	90.88%	9.12%	0.0%
Small-World #5 Degree $\langle k \rangle = 32$ Nodes=100	1600	1%	96.83%	75.0%	2.25%	97.75%	0.0%
		2%	95.49%	74.5%	2.81%	97.19%	0.0%
		4%	89.12%	72.5%	7.69%	92.31%	0.0%
		8%	77.88%	63.75%	18.13%	81.87%	0.0%
		16%	64.52%	55.0%	31.0%	69.0%	0.0%
		32%	53.38%	50.0%	68.38%	31.63%	0.0%
		64%	45.78%	50.0%	93.5%	6.5%	0.0%
Small-World #6 Degree $\langle k \rangle = 64$ Nodes=100	3200	1%	99.23	100.0%	33.72%	66.28%	0.0%
		2%	98.82	100.0%	33.53%	66.47%	0.0%
		4%	97.32	100.0%	36.69%	63.31%	0.0%
		8%	95.17	100.0%	43.5%	56.5%	0.0%
		16%	92.57	100.0%	55.16%	44.84%	0.0%
		32%	89.27	100.0%	84.47%	15.53%	0.0%
		64%	87.85	100.0%	96.19%	3.81%	0.0%

圖 47 到圖 52 也是 WS 小世界網路的實驗結果。我們固定繞線機率 p ，觀察平均節點分支度 $\langle k \rangle$ 的變化與連結種類分佈之間的關係。我們發現當繞線機率 p 較小時，強鍵式連結佔了所有連結大部分的比例，隨著 p 增加，橋接式連結的比例也跟著增加，這樣的結果和章節 4.1.1 中的驗證結果相同。此外，平均分支度為 2 時，整體網路為直鏈狀網路，並且經過繞線之後很可能成為不連通圖，因此才會出現沉沒橋接式連結。而當平均分支度為 64 時，整體網路的連結相當密集，平均分隔度低，因此許多連結被判斷成橋接式連結。一般來說，不同分支度對於橋接式連結與強鍵式連結的分佈狀況並沒有直接的影響。

1. 節點個數：100、繞線機率：2%。

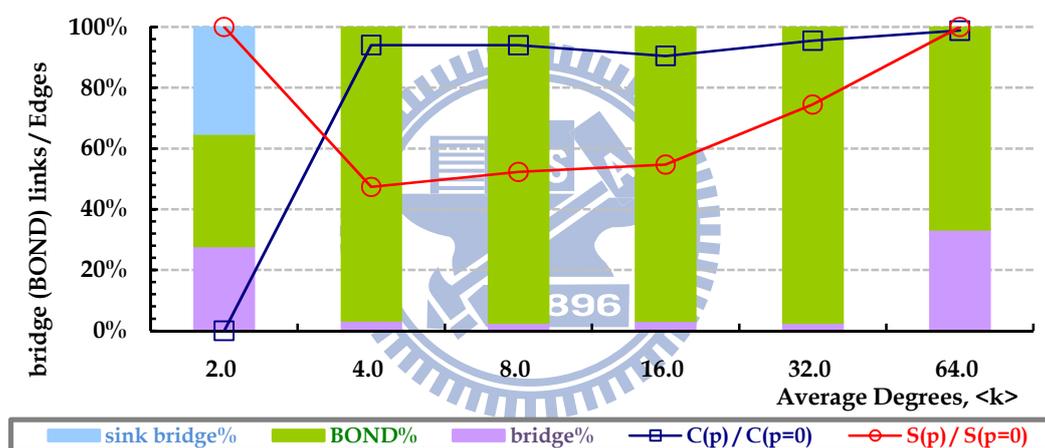


圖 47：WS-small world, $N=100$, $p=2\%$ 。

2. 節點個數：100、繞線機率：4%。

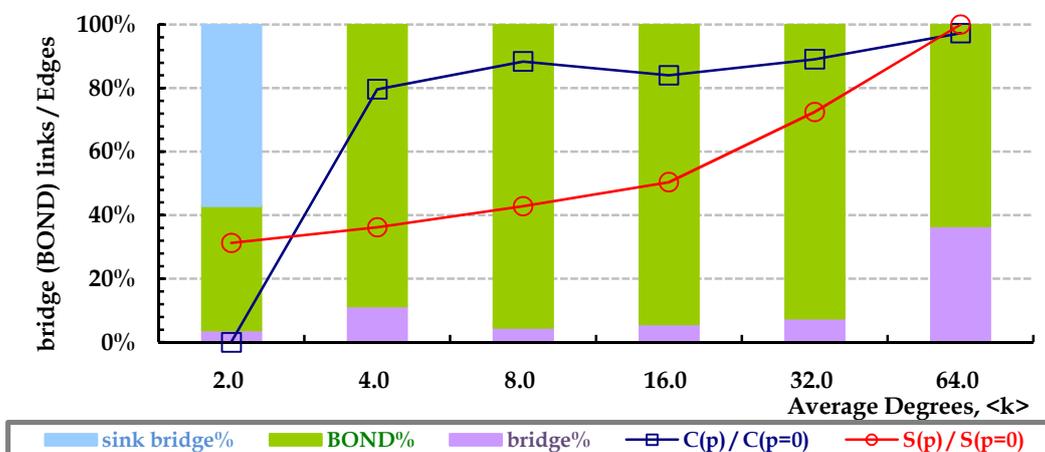


圖 48：WS-small world, $N=100$, $p=4\%$

3. 節點個數：100、繞線機率：8%。

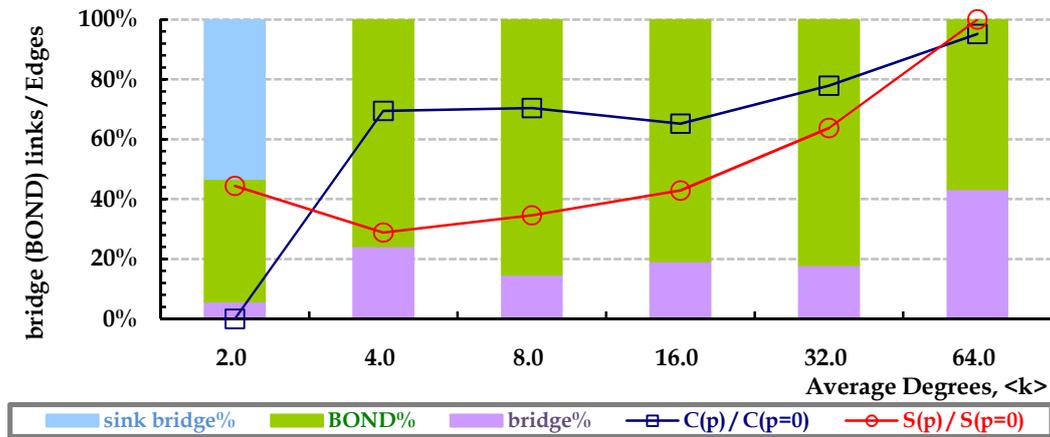


圖 49：WS-small world, N=100, p=8%。

4. 節點個數：100、繞線機率：16%。

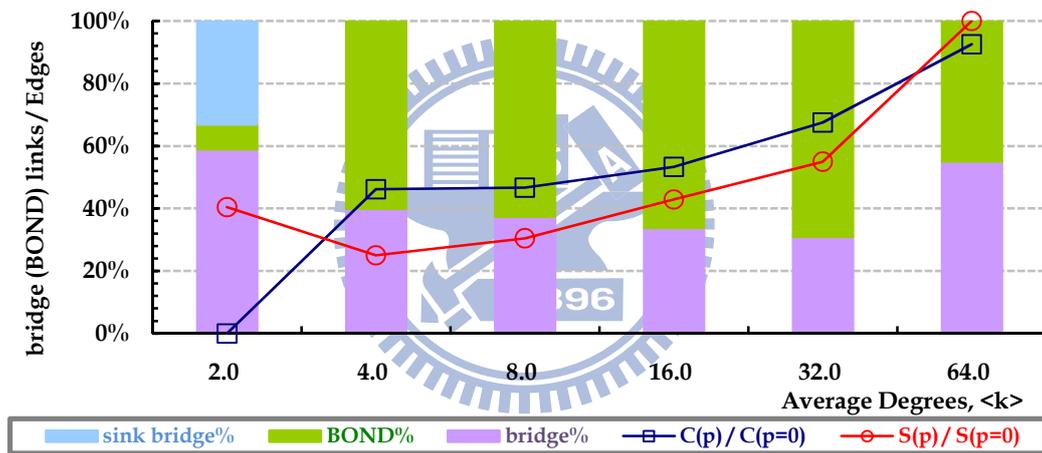


圖 50：WS-small world, N=100, p=16%。

5. 節點個數：100、繞線機率：32%。

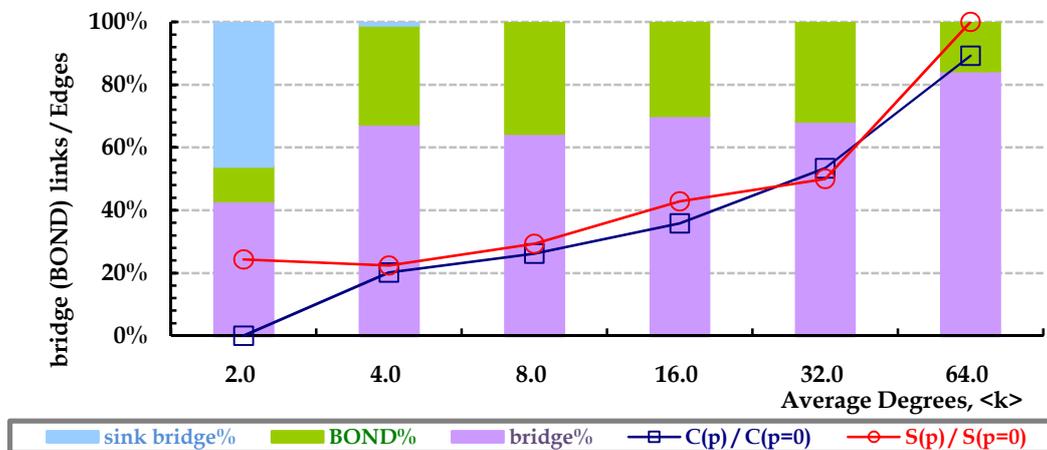


圖 51：WS-small world, N=100, p=32%。

6. 節點個數：100、繞線機率：64%。

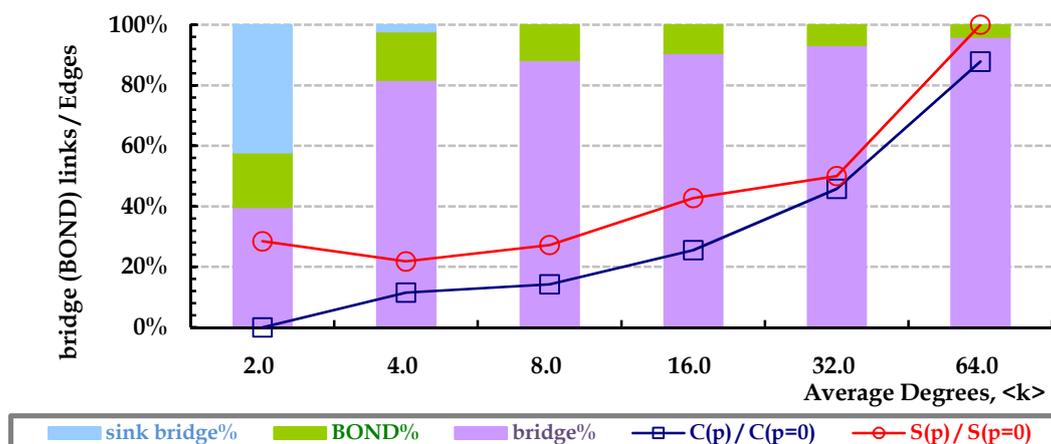


圖 52：WS-small world, N=100, p=64%。

NW 小世界網路

接下來利用 Newman 和 Watts 所提出的方法產生出一連串的小世界網路。節點個數為 100，固定機率 p 在隨機選取的一對節點之間加上一條連結，觀察節點平均分支度 $\langle k \rangle$ 的變化與橋接式連結和強鍵式連結的分佈之間的關係 ($\langle k \rangle = 2, 4, 8, 16, 32, 64$)。下列各圖中紫色三角形所構成的曲線代表連結密度，其計算方式如下：

$$Density = \frac{2 \times E}{N \times (N - 1)}$$

其中 E 為網路中的連結個數， N 為網路中的節點個數。圖 53 到圖 58 是實驗結果圖，我們可以發現到當 p 增加時，網路中的橋接式連結增加的幅度沒有 Watts 和 Strogatz 所產生的小世界網路所增加的幅度大。這是因為 Newman 和 Watts 所產生的小世界網路並不是隨機繞線而是隨機加連結，而原本的網路為規則網路，被加入的連結可能也會被判斷成為強鍵式連結，因此橋接式連結的增加幅度會較為緩慢，這個現象也接連影響到網路群聚度並沒有太大的改變。表 6 為 NW 小世界網路的統計數據。

1. 節點個數：100、隨機加邊機率 p ：2%。

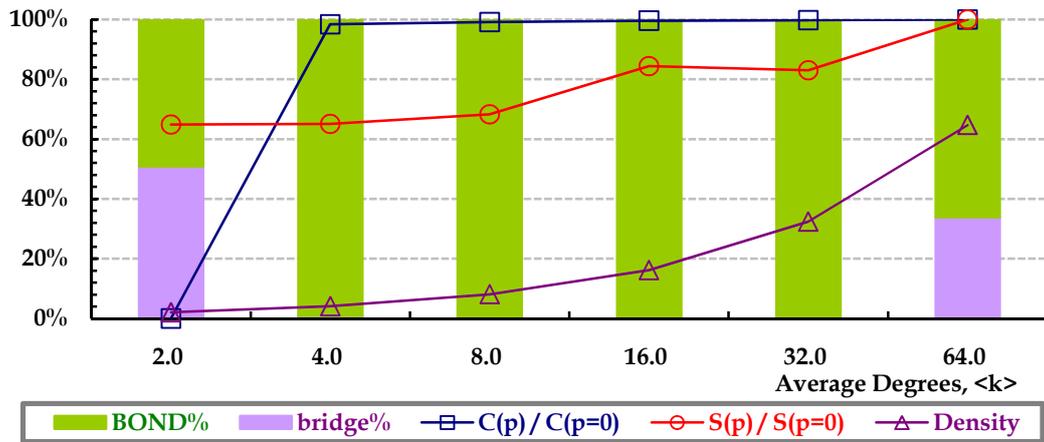


圖 53：NW-small world, $N=100$, $p=2\%$ 。

2. 節點個數：100、隨機加邊機率 p ：4%。

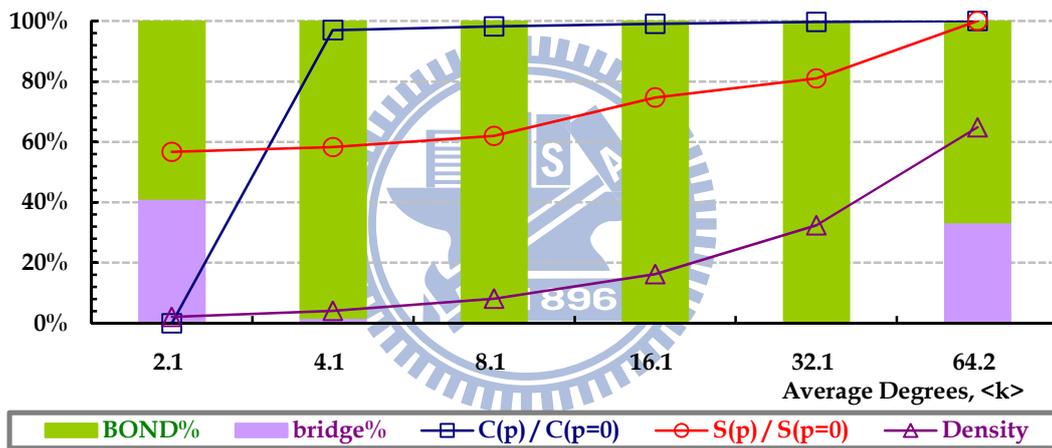


圖 54：NW-small world, $N=100$, $p=4\%$ 。

3. 節點個數：100、隨機加邊機率 p ：8%。

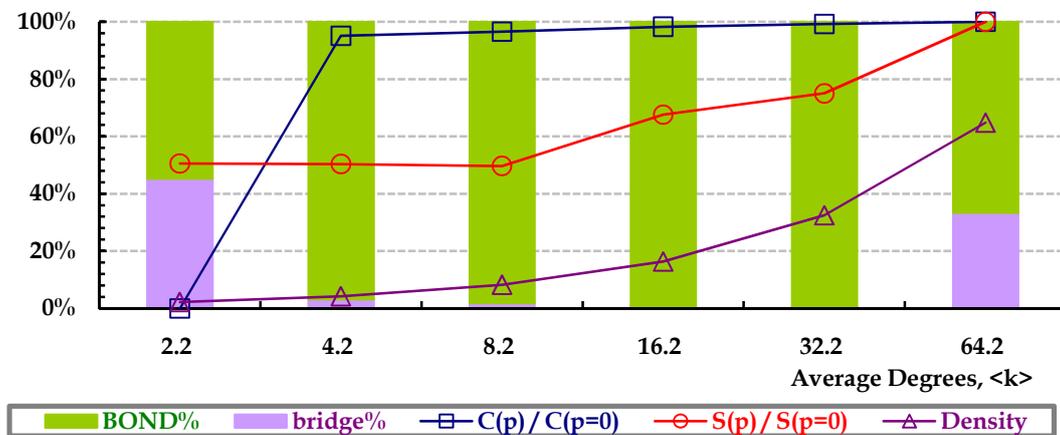


圖 55：NW-small world, $N=100$, $p=8\%$ 。

4. 節點個數：100、隨機加邊機率 p ：16%。

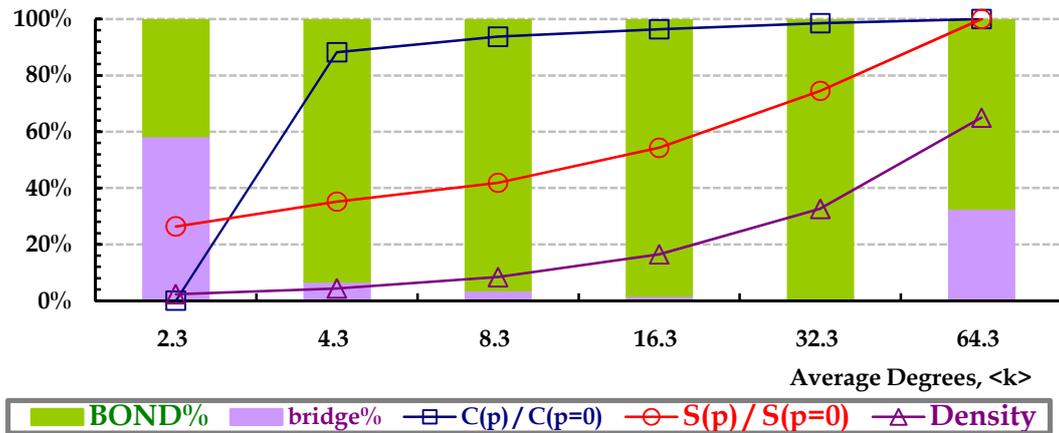


圖 56：NW-small world, N=100, $p=16\%$ 。

5. 節點個數：100、隨機加邊機率 p ：32%。

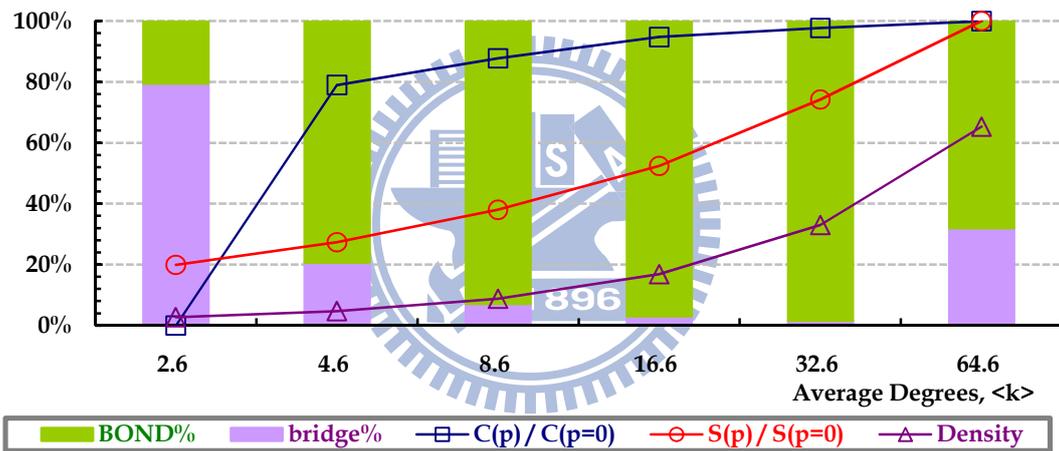


圖 57：NW-small world, N=100, $p=32\%$ 。

6. 節點個數：100、隨機加邊機率 p ：64%。

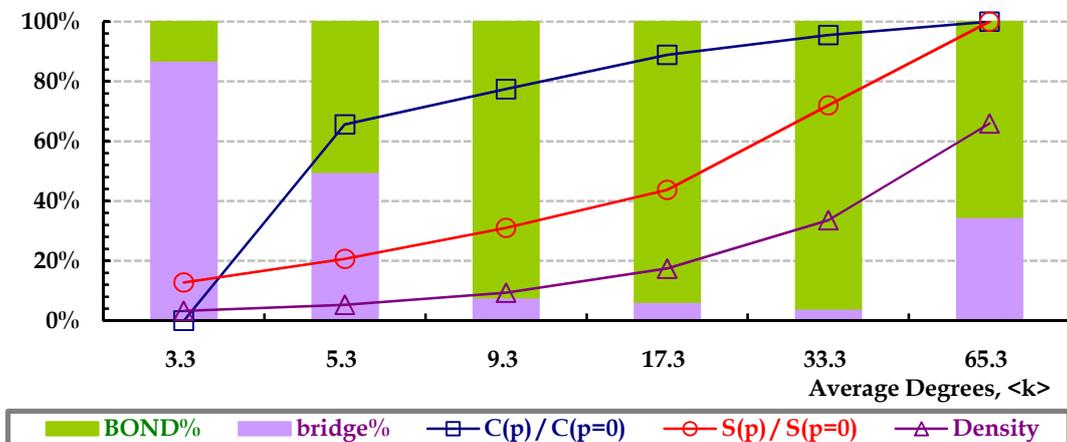


圖 58：NW-small world, N=100, $p=64\%$ 。

表 6：不同繞線機率下的 NW 小世界網路

Category	Edges	Rewiring(p)	$C(p)/C(p=0)$	$S(p)/S(p=0)$	Bridge link	Bond link	Sink link
Small-World #1 Degree $\langle k \rangle=2$ Nodes=100	100	1%	0.0%	92.64%	91.0%	9%	0.0%
		2%	0.0%	64.88%	50.98%	49.02%	0.0%
		4%	0.0%	56.70%	41.35%	58.65%	0.0%
		8%	0.0%	50.56%	45.37%	54.63%	0.0%
		16%	0.0%	26.38%	58.62%	41.38%	0.0%
		32%	0.0%	19.92%	79.55%	20.45%	0.0%
		64%	0.0%	12.8%	87.20%	12.8%	0.0%
Small-World #2 Degree $\langle k \rangle=4$ Nodes=100	200	1%	99.2%	80.16%	0.5%	99.5%	0.0%
		2%	98.4%	65.08%	0.99%	99.01%	0.0%
		4%	97.0%	58.28%	2%	98%	0.0%
		8%	95.13%	50.36%	3.37%	96.63%	0.0%
		16%	88.2%	35.16%	6.94%	93.06%	0.0%
		32%	78.98%	27.32%	20.69%	79.31%	0.0%
		64%	65.62%	20.68%	49.91%	59.09%	0.0%
Small-World #3 Degree $\langle k \rangle=8$ Nodes=100	400	1%	99.55%	80.92%	0.2%	99.8%	0.0%
		2%	99.11%	68.31%	0.5%	99.5%	0.0%
		4%	98.22%	62.0%	0.99%	99.01%	0.0%
		8%	96.51%	49.69%	1.96%	98.04%	0.0%
		16%	93.73%	41.85%	3.85%	96.15%	0.0%
		32%	87.77%	38.0%	7.18%	92.82%	0.0%
		64%	77.41%	31.0%	7.97%	92.03%	0.0%
Small-World #4 Degree $\langle k \rangle=16$ Nodes=100	800	1%	99.76%	89.14%	0.1%	99.9%	0.0%
		2%	99.60%	84.43%	0.25%	99.75%	0.0%
		4%	99.06%	74.71%	0.5%	99.5%	0.0%
		8%	98.27%	67.57%	0.99%	99.01%	0.0%
		16%	96.39%	54.29%	1.96%	98.04%	0.0%
		32%	94.74%	52.43%	3.13%	96.87%	0.0%
		64%	88.86%	43.71%	6.48%	93.52%	0.0%
Small-World #5 Degree $\langle k \rangle=32$ Nodes=100	1600	1%	99.88%	87.5%	0.06%	99.94%	0.0%
		2%	99.76%	83.0%	0.12%	99.88%	0.0%
		4%	99.70%	81.0%	0.25%	99.75%	0.0%
		8%	99.24%	75.0%	0.5%	99.05%	0.0%
		16%	98.51%	74.5%	0.93%	99.07%	0.0%
		32%	97.71%	74.25%	1.65%	98.35%	0.0%
		64%	95.51%	72.0%	4.27%	95.73%	0.0%
Small-World #6 Degree $\langle k \rangle=64$ Nodes=100	3200	1%	100.0%	100.0%	34.11%	65.89%	0.0%
		2%	100.0%	100.0%	33.98%	66.02%	0.0%
		4%	100.0%	100.0%	33.55%	66.45%	0.0%
		8%	100.0%	100.0%	33.54%	66.46%	0.0%
		16%	100.0%	100.0%	32.87%	67.13%	0.0%
		32%	100.0%	100.0%	32.02%	67.98%	0.0%
		64%	100.0%	100.0%	34.87%	65.13%	0.0%

附錄 F. 資料來源

表 F.1：本研究所有真實網路資料來源

網路種類	網路名稱	節點數	連結數	來源
社會網路	成衣工廠員工關係網路	14	28	(Kilduff & Tsai, 2003)
	夏令營參加者關係網路	18	35	
	南部村落婦女關係網路	18	50	(Freeman, 2004)
	Zachary 空手道網路	34	78	(Girvan & Newman, 2002)
	囚犯朋友網路	67	182	(Milo et al., 2004)
	領導者課程網路	32	96	
食物鏈網路	Little Rock	92	984	(Huang et al., 2007)
	Ythan	83	391	
	St. Martin	42	205	
	Chesapeake	31	67	
	Coachella	29	243	
	Skipwith	25	189	
	B. Brook	25	104	
電子網路	S208	122	189	(Huang et al., 2007)
	S420	252	399	
	S838	512	819	

参考文献

- Albert, R. and Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47-97.
- Artzy-Randrup, Y., Fleishman, S. J., Ben-Tal, N., and Stone, L. (2004). Comment on "network motifs: Simple building blocks of complex networks" and superfamilies of evolved and designed networks". *Science*, 305(5687):1107c+.
- Barabasi, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509-512.
- Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2003). The architecture of complex weighted networks.
- Bascompte, J., Melián, C. J., and Sala, E. (2005). Interaction strength combinations and the overfishing of a marine food web. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15):5443-5447.
- Berlow, E. L. (1999). Strong effects of weak interactions in ecological communities. *Nature*, 398(6725):330-334.
- Berkman, L. F. and Syme, S. L. (1979). Social networks, host resistance, and mortality: a nine-year follow-up study of alameda county residents. *Am J Epidemiol*, 109(2):186-204.
- Brglez, F., Bryan D. and Kozminski, K. (1989). Combinational Profile of Sequential Benchmark Circuits. *In Proc. Of the Int. Symp on Circuits and Systems*, pp. 1929-1934.
- Cancho, R. F., Janssen, C., and Solé, R. V. (2001). Topology of technology graphs: small world patterns in electronic circuits. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 64(4 Pt 2).
- Clauset, A., Moore, C., and Newman, M. E. J. (2008). Hierarchical structure

and the prediction of missing links in networks. *Nature*, 453(7191):98-101.

Csermely, P. (2006). *Weak Links*. Springer-Verlag .

Dekker, A. (2007). Studying Organizational Topology with Simple Computational Models. *Journal of Artificial Societies and Social Simulation*, vol. 10, no. 4 6.

Dunne, A.J., Williams, R.J. and Martinez, N.D. (2002). Small network but not small world: unique aspects of food web structure. *Santa Fe Institute*, 02-03-010.

Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215-239.

Freeman, L. (2004). Finding Groups: A Meta-Analysis of Southern Women Data.

Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821-7826.

Granovetter, M.S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360-1380.

Goh, K.-I., Oh, E., Jeong, H., Kahng, B., and Kim, D. (2002). Classification of scale-free networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12583-12588.

Goyal, S. (2008). *An introduction to the Economics of Networks*. Princeton University Press.

Halpern, D. (2008). *Social Capital*. Polity Press.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44(1):82-111.

Huang, C.Y., Cheng, C.Y. and Sun, C.T. (2007). Bridge and Brick Network Motifs: Identifying Significant Building Blocks from Complex Biological Systems. *Artificial Intelligence in Medicine*, 41(2), 117-127.

Kilduff, M. & Tsai, W. (2003). *Social Network and Organizations*. Thousand Oaks, CA: Sage.

Mccann, K., Hastings, A., and Huxel, G. R. (1998). Weak trophic interactions and the balance of nature. *Nature*, 395(6704):794-798.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824-827.

Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U. (2004). Superfamilies of evolved and designed networks. *Science*, 303(5663):1538-1542.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167-256.

Onnela, J. P., Saramaki, J., Hyvonen, J., Szabo, G., Lazer, D., Kaski, K., Kertesz, J., and Barabasi, A. L. (2006). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332-7336.

Papakyriazis, N. and Boudourides, M. (2001). Electronic weak ties in network organisations. In *4th GOR Conference*, Goettingen, Germany.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., and Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2658-2663.

Stokman, F.N., Ziegler, R., Scott, J. (1985). *Networks of Corporate Power: A Comparative Analysis Of Ten Countries*. Polity Press.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268-276.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, Cambridge University Press.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440-442.

Williams, R. J. and Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774):180-183.

