複雜群落網路中訊息傳播的數學分析與模擬

蔡瓊萩

南亞技術學院資訊工程系

摘要

在本論文中，我們將借助一個兩狀態的 SIS 疾病傳染模型研究複雜群落網路中訊息傳播的動力行為。這裡所謂的群落網路乃指具有某種群落結構的一種社會網路，在這種群落網路中，整體網路可區分成數個子群，同一子群內的個體間聯結密切，而不同子群間個體的聯結則相對地較為稀疏，網路中的每一個節點代表一個個體，線段則代表個體之間具有直接的聯結與相互作用。在本論文中，我們首先將建構一個靜態的群落結構網路,進而研究在此網路中訊息擴散的動力行為，特別地，我們將研究網路結構例如群落度和群聚度等要素如何影響複雜群落網路中的訊息傳播行為，同時亦將透過數學程式語言 MATLAB 進行相關的數值模擬實驗。

關鍵詞：訊息傳播，複雜網路，群落度，群聚度，傳染病模型，穩定性

Email： cctsai@nanya.edu.tw　Tel：03-4361070#6943

ANALYSIS AND SIMULATION OF INFORMATION PROPAGATION IN COMPLEX MODULAR NETWORKS

Chiung-Chiou Tsai

Department of Computer Science and Information Engineering
Nanya Institute of Technology
Jhongli, Taiwan

Abstract

In this paper, we will study the dynamics of information propagation in complex modular networks by using a two-state SIS epidemic model. Here a modular network means a social network with the community structure. In such a modular network, there are groups of nodes with many connections between their members and few connections to nodes outside the group, in which the nodes of the network represent individuals and the edges represent various interactions among those individuals. In this paper, we will first construct a static model of modular network and then investigate the dynamical behavior of the information propagation. In particular, we will study how the network structure such as the community degree and the clustering coefficient affect the information propagation in the complex modular network. Numerical experiments will be performed by using the MATLAB programming language.

Keywords: information propagation, complex network, community degree, clustering coefficient, epidemic model, stability

Email： cctsai@nanya.edu.tw　Tel：03-4361070 ext.6943

壹、前言

最近十餘年來，複雜網路(complex network)的研究是一個熱門的跨領域主題，它包含了數學、物理、生物、資訊和社會科學等多種豐富的知識內容，可參考 Newman 的著作[20]及其書中大量的參考文獻。基本上，對複雜網路的研究大致上可分成兩個大方向，一為網路的拓撲結構(topology structure)，另一為網路上訊息的傳遞與擴散(information propagation)。

關於複雜網路的拓撲結構研究，以數學的角度來看，網路可視為一個包含點和邊(即線段)的圖形，屬於離散數學中的圖形理論(graph theory)領域，但其拓撲研究重點在於網路中任兩點之間的分隔度(degree of separation)、連結度分布(degree distribution)、群聚度(clustering coefficient)和各種點和邊的中心度(centrality)等網路結構特徵，因此一般依據分隔度和群聚度這兩種特性引進了所謂的小世界網路(small-world network)理論。小世界網路可視為一種介在一般古典的隨機網路(random network)和規則網路(regular network)之間的特別結構。目前文獻上最常見的兩類型小世界網路為一種不具極高連結數的點的 Newman-Watts-type 小世界網路和具極高連結數的點的 Barabasi-Albert 無尺度網路(其 degree 的分佈滿足所謂的冪次率)。

第二種複雜網路研究方向主要探討網路上訊息傳遞與擴散，探討在一給定結構的複雜網路上訊息包如何的傳遞、流言如何擴散、病毒如何傳染、神經元間如何同步運作等，同時討論其和網路結構性質的關聯性，當然在此網路上數學模型的建立是關鍵主題之一。神經元之間的同步研究一般是設定成在一給定的網路上，將每一點視為一個神經元(neuron)，其動力行為被一組常微分方程所描述，例如 FitzHugh-Nagumo neuron, Hindmarsh-Rose neuron, 或是人工的 Chua's circuit 等，再經由合適的耦合形成一個大型複雜網路動力系統(請參考[11])。而訊息傳遞擴散的數學模型一般就以傳染病模型來代替，將訊息視為某種病毒，這應該是一相當合理的想像，這也是本論文所將採用的方式。

最近幾年來，關於社會網路已經有廣泛的探討，研究結果指出社會網路這種小世界網路通常具有某種群落結構(community structure)，具體而言，整體網路可區分成數個子群，同一子群內的個體間聯結密切，而不同子群間個體的聯結則相當稀疏。直觀上，我們可以想像在真實世界中如流言擴散、傳染病擴散等應該是發生在具群落結構的社會網路中，網路中的節點代表一個個體，線段則代表個體之間的具有直接的聯結與相互影響。一個理想的群落結構網路可想像如圖 1，而

具備小世界特性的群落結構網路可經由適當的加入捷徑(紅色線段)達成(請參考[11])。圖 1 呈現一個理想中的群落結構網路，分布在一個封閉環上，它具有 105 個點並且分成 21 個子群，同一子群內的個體相互聯結，每一子群僅與左右最近的相鄰子群有一聯結。加入紅色線段捷徑形成一小世界，此處我們採用 Newman-Watts 的方法，任兩子群中的任兩個點以極小的機率決定是否加上新的聯結捷徑(紅色線段)。

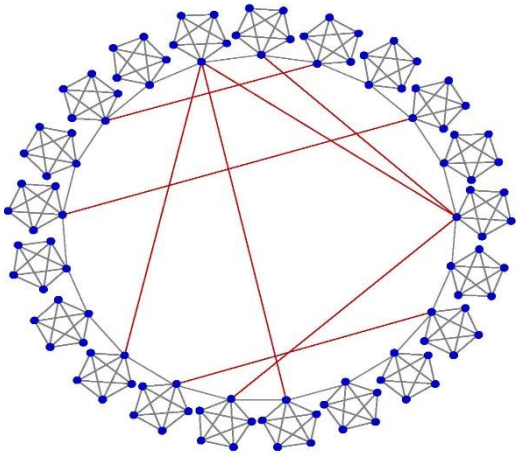

圖 1: A Newman-Watts-type small-world network with community structure [11]

當然大型群落網路的建構亦可由無尺度網路出發，連結幾個無尺度網路形成一個大型具群落結構的無尺度網路 (請參考[11])。

貳、文獻探討

上述提及我們將以傳染病模型來當作訊息傳遞擴散的數學模型，下面我們簡述疾病傳染數學模型的相關發展。有關以微分方程來描述疾病傳染動態演化最早的數學模型應該是開始於 1970 年代中末期由 Anderson-May [1] 和 Hethcote [6] 所提出的常微分方程組，歷經三十餘年來已經引起相當廣泛的研究興趣並取得相當的進展。典型的傳染病模型(例如 SIR 模型)乃描述某種特別疾病易受感染者之個體數、已染病之個體數及染病後復原之個體數的動態變化。在 SIR 傳染病模型中，驅使疾病傳染擴散的力量來自易受感染個體變成染病個體的一個比率變數，一般而言這個比率變數是已染病個體數的函數，此變數同時包含個體間交互影響而造成傳染病擴散流行的某些訊息。在進一步的適當假設下，這個比率變數將衍生出一個含有易受感染之個體和已染病之個體間接觸比率參數，此接觸比率參數和傳染病的傳染門檻有密切關係，同時亦將使得傳染病模型具有相當豐富的動態系統

行為。

假設 $S(t), I(t), R(t)$ 分別代表易受感染個體數、已被感染個體數及已復原個體數，則此 SIR 模型可以下列常微分方程組描述：

$$\begin{cases} S'(t) = -\beta S(t)I(t) - \mu S(t) + \mu N(t), \\ I'(t) = \beta S(t)I(t) - \mu I(t) - \lambda I(t), \\ R'(t) = \lambda I(t) - \mu R(t), \end{cases} \tag{1}$$

其中 $S(t) + I(t) + R(t) = N(t)$ 為個體總數，而 $\beta, \mu, \lambda > 0$ 為常數，其中 β 即所謂的個體間接觸比率參數。在這個模型中，個體總數被假設為一正常數，同時假設出生與死亡率相等，若進一步假設 $N(t) = 1,\ t \geq 0$，則此時 $S(t), I(t), R(t)$ 可分別視為總個體數中易受感染者之個體數、已染病之個體數及染病後復原之個體數的比率數。

在 70 年代中末期，Anderson-May [1] 和 Hethcote [6] 分別證明了

$$\delta \equiv \frac{\beta}{\lambda + \mu} > 1$$

是傳染病發生的一個門檻，以生物及數學語言來描述，若 $\delta \leq 1$ 則該疾病終將消失，即邊界平衡點 $E_0 = (1,0,0)$ 為全局漸近穩定；若 $\delta > 1$ 則該疾病將持續存在而趨近於另一平衡點

$$E_+ = (\frac{1}{\delta}, \frac{\mu(\delta - 1)}{\beta}, \frac{\lambda(\delta - 1)}{\beta})$$

即內部平衡點 E_+ 為全局漸近穩定。

在過去的三十餘年間，陸續有眾多日、中、歐美學者進一步研究類似的傳染病傳播的數學模型或其變形，特別是引入了時間遲滯的概念，例如上述方程系統(1)中 $S'(t)$ 和 $I'(t)$ 分別受到 $\beta S(t)I(t)$ 的影響，其中具傳染力的 $I(t)$ 應不會即時呈現出來，而會有時間遲滯現象變為 $\beta S(t)I(t-h)$，其中 $h > 0$ 常數。例如在 2004 年間 Ma-Song-Takeuchi [16] 等人研究下列具時間遲滯的系統：

$$\begin{cases} S'(t) = -\beta S(t)I(t-h) - \mu_1 S(t) + b, \\ I'(t) = \beta S(t)I(t-h) - \mu_2 I(t) - \lambda I(t), \\ R'(t) = \lambda I(t) - \mu_3 R(t). \end{cases} \tag{2}$$

他們證明了在假設條件

$$\frac{b}{\mu_1} > \frac{\mu_2 + \lambda}{\beta} \equiv S^*$$

的情況下，對任意的 $h > 0$ 系統(2)具有永久性(permanence)，且當 $h > 0$ 充分小時，

其內部平衡點

$$E_{+}=(S^{*},\frac{(b-\mu_1 S^{*})}{\beta S^{*}},\frac{\lambda(b-\mu_1 S^{*})}{\mu_3\beta S^{*}})$$

為全局漸近穩定。

一般而言，時間遲滯除了上述離散式時間遲滯外，另一種更為接近真實的是分佈式時間遲滯。例如 Beretta et al. [4] 和 Ma et al. [17] 等人分別於 2001 及 2002 年提出下列分佈式時間遲滯模型：

$$\begin{cases} S'(t)=-\beta S(t)\int_0^h I(t-s)d\eta(s)-\mu_1 S(t)+b, \\ I'(t)=\beta S(t)\int_0^h I(t-s)d\eta(s)-\mu_2 I(t)-\lambda I(t), \\ R'(t)=\lambda I(t)-\mu_3 R(t), \end{cases} \tag{3}$$

其中 $\eta(s):[0,h]\to R$, 非遞減函數、有界變動且 $\int_0^h d\eta(s)=\eta(h)-\eta(0)=1$。儘管考慮的時間遲滯方式不同，他們仍然得到和系統(2)極為類似的結論。除了上述幾個常見模型外，還有相當多的類似變形。

另一方面，在我們前述的 SIR 傳染病模型中，傳染病的傳染門檻和接觸比率參數 β 有密切而直接的關係，

$$\delta\equiv\frac{\beta}{\lambda+\mu}>1\text{。}$$

顯而易見接觸比率參數 β 將和網路的拓撲結構有直接的關聯。因此，在 2001 和 2002 年間 Pastor-Satorras 和 Vespignani 提出了一個著名的 SIS 模型描述如下(參見[21, 22])：

$$\frac{d\rho_k(t)}{dt}=-\rho_k(t)+\lambda(k)(1-\rho_k(t))\Theta(t)\quad for\ \ k=1,2,\cdots,n. \tag{4}$$

其中 $\rho_k(t)$ 是 degree 為 k 的已被感染個體數的密度，也就是 $1-\rho_k(t)$ 是 degree 為 k 的未被感染個體數的密度，$\lambda(k)$ 是疾病傳染比率。$\Theta(t)$ 是任給定一個邊而該邊會連到已經被感染個體數的比率，其數學關係式可表示為

$$\Theta(t)=\frac{\sum_{i=1}^{n} iP(i)\rho_i(t)}{<k>},$$

其中 $P(i)$ 是 degree 為 i 的點的機率，而 $<k>=\sum_{i=1}^{n} iP(i)$ 則是整個網路的平均 degree (mean degree)。Pastor-Satorras 和 Vespignani 的工作，成功地把疾病傳染和

複雜網路的結構連結在一起，獲得極大的回響，進而有大量不同模型的相關研究成果出現。

參、研究方法

在本論文中，我們主要研究群落結構如何影響某一 SIS 傳染病模型在社會網路中個體間的訊息傳播(疾病傳染)動態行為。我們將從建構一個靜態的群落結構網路開始，進而研究在此網複雜路中訊息的擴散情況，同時亦將透過數學程式語言 MATLAB 進行相關的數值模擬實驗。

我們以某個 SIS 模型為例[26]，研究群落結構如何影響 SIS 在社會網路中個體之間的訊息傳遞擴散。首先我們將建構一個靜態的群落結構網路，進而研究在此網路中訊息的擴散情況。關於此靜態的群落結構網路的建構初步構想如下(參見[15])：

1. 考慮一個具有 N_0 個節點的網路，隨機分配成 m 個子群，每一子群有 $n_i\ (i=1,2,\cdots,m)$ 個節點。因此，

$$\sum_{i=1}^{m} n_i = N_0 \text{ 。}$$

2. 在每一子群內，我們以機率 p 去聯結任兩個節點。因此，在子群 i 內，我們共有 $\frac{1}{2}n_i(n_i-1)p$ 條線段。
3. 在任兩個子群間，我們以機率 q 去聯結它們之間的任兩個節點。因此，在子群 i 和子群 j 間，我們共有 $n_i n_j q$ 條線段。

此群落結構網路的外觀概略如下示意圖[15]：

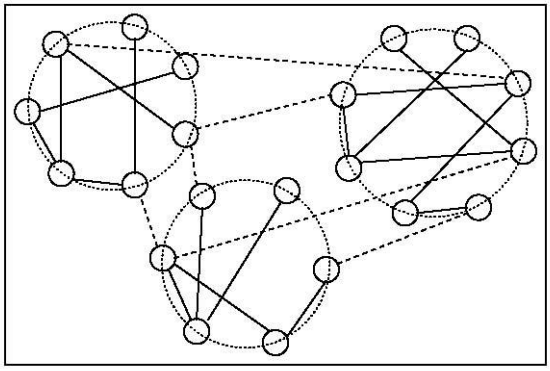

圖 2：群落結構網路示意圖

若假設該群落結構網路的總線段數是 N, 則我們有

$$N=\sum_{i=1}^{m}\frac{1}{2}n_i(n_i-1)p+\sum_{i<j}^{m}n_in_jq \text{ 。}$$

進一步我們令$\sigma=\dfrac{p}{q}$，則調整σ值的大小，將決定此群落結構網路的群落度(degree of community)強弱，例如$\sigma=1$時，此網路不具群落結構而是一種隨機網路；當$\sigma>>1$時則同一子群內的個體間聯結密切，而不同子群間個體的聯結則相對稀疏，所以是一個具較強群落結構的網路。

我們將研究在此類的群落結構網路上 SIS 訊息傳遞擴散的情況，探討群落結構如何影響訊息傳遞擴散在社會網路中個體間的動態行為。所考慮的 SIS 模型如下(參見[26])：

$$\begin{cases}\dfrac{dS_k(t)}{dt}=b(S_k(t)+I_k(t))-dS_k(t)-\lambda(k)S_k(t)\Theta(t)+aI_k(t),\\ \dfrac{dI_k(t)}{dt}=\lambda(k)S_k(t)\Theta(t)-(a+d)I_k(t),\end{cases} \tag{5}$$

$k=1,2,\cdots,n.$此模型考慮了出生率及死亡率的影響，這裡的S_k與I_k分別代表 degree 為k的受感染個體數和已被感染個體數的密度，$\lambda(k)>0$是疾病傳染的比率，b 是指出生率，d是死亡率，而a是復原率。$\Theta(t)$描述網路中任給定一個邊而該邊會連到已被感染個體數的比率，其數學關係式可表示為：

$$\Theta(t)=\frac{\sum_{i=1}^{n}\varphi(i)P(i)I_i(t)}{<k>},$$

其中$\varphi(i)$指的是i的連邊裡會傳染疾病的狀態，$\varphi(i)$一般可以是一個非線性函數(參見 [26])。

肆、結果與討論

我們對於訊息傳遞擴散最感興趣的問題是介在訊息擴散與不擴散間，有沒有一個引爆點(epidemic threshold)，只要跨過這條分隔線時，訊息(疾病)就會傳染開來。相關文獻已經指出模型(4)的確具有一個 epidemic threshold 為

$$\lambda_c=\frac{<k>}{<k^2>}$$

當$\lambda\geq\lambda_c$時，訊息會傳染開來且會持續存在；反之則訊息會消失。這個結論的嚴謹數學證明可參考 Wang 和 Dai [23]。

我們的群落結構網路上 SIS 訊息傳遞擴散模型考慮了出生率及死亡率的影響：

$$\begin{cases} \dfrac{dS_k(t)}{dt} = b(S_k(t)+I_k(t)) - dS_k(t) - \lambda(k)S_k(t)\Theta(t) + aI_k(t), \\ \dfrac{dI_k(t)}{dt} = \lambda(k)S_k(t)\Theta(t) - (a+d)I_k(t), \end{cases} \quad k = 1,2,\cdots,n \qquad (6)$$

這裡的 S_k 與 I_k 分別是 degree 為 k 的 susceptible node 與 infected node 的 density；$\lambda(k)>0$ 是(k-dependent) infection rate；b 是出生率；d 是死亡率；a 是復原率；$\Theta(t)$ 要描述的是網路中訊息會傳遞擴散的邊的比例：

$$\Theta(t) = <k>^{-1} \sum_{i=1}^{n} \varphi(i)P(i)I_i(t)$$

其中 $\varphi(i)$ 指的是 i 個連邊裡訊息會傳遞擴散的狀態(infectivity)，$\varphi(i)$ 為一個非線性函數。我們可以驗證 threshold value 為(參見 Zhu et al. [26])

$$R_0 = \frac{<\lambda(k)\varphi(k)>}{(a+b)<k>},$$

其中 $<\lambda(k)\varphi(k)> = \sum_{i=1}^{n} \lambda(i)\varphi(i)P(i)$。

當 $R_0>1$ 時，亦即是訊息傳遞擴散會持續存在。相反地，當 $R_0 \le 1$ 時，訊息傳遞擴散最終會消失。而且出生率也會影響這個 threshold value，出生率越高，R_0 越小訊息傳遞擴散越不容易爆發。

我們已經完成部份 system (5) 的 global dynamics 數值實驗，系統的動態行為如下圖：

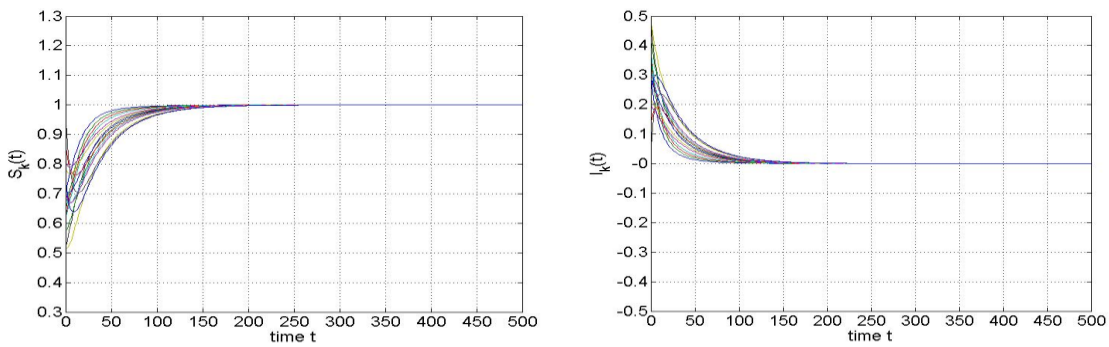

圖 3：系統(5)的動態行為

當調整 system (5) 裡的參數 λ 以後，我們發現其動態行為改變如下圖：

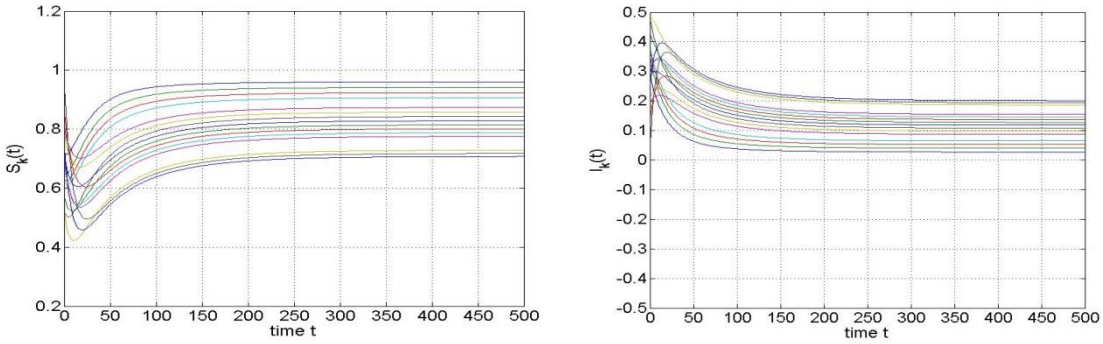

圖 4：改變系統(5) 參數 λ 後的動態行為

我們可以從上述的數值模擬發現，確實會有一個 threshold value 把這兩種動態行為分隔開來。第一組圖形對應的是解趨於 disease-free equilibrium，第二組圖形則對應的是解會趨近於 endemic equilibrium。另外從數值實驗上的觀察我們可發現這兩個平衡點是 globally asymptotically stable，因此研究在什麼條件下，endemic equilibrium 會存在且是 globally asymptotically stable，將是我們下一步的研究主題，而建構出一個適當的 Lyapunov function 將會是最大的難題。

伍、結論

在此論文中，我們藉由一個兩狀態的SIS疾病傳染模型研究複雜群落網路中訊息傳播的動力行為。所謂的群落網路乃指具有某種群落結構的一種社會網路，在這種群落網路中，整體網路可區分成數個子群，同一子群內的個體間聯結密切，而不同子群間個體的聯結則相對地較為稀疏，網路中的每一個節點代表一個個體，線段則代表個體之間具有直接的聯結與相互作用。我們驗證了訊息傳遞擴散的問題存在一個引爆點(epidemic threshold)，

$$R_0 = \frac{<\lambda(k)\varphi(k)>}{(a+b)<k>} \text{ 。}$$

當 $R_0 > 1$ 時，亦即是訊息傳遞擴散會持續存在。相反地，當 $R_0 \le 1$ 時，訊息傳遞擴散最終會消失。而且出生率也會影響這個引爆值，顯然如果出生率越高，則 R_0 值越小，訊息傳遞擴散越不容易爆發。同時我們透過數學程式語言MATLAB進行了相關的數值模擬實驗，實驗結果符合數學理論分析。

陸、參考文獻

[1] R. M., Anderson and R. M. May, 1979, “Population biology of infectious diseases:

part I." *Nature*, 280: pp. 361-367.

[2] S. Bansal, B. T. Grenfell, and L. A. Meyers, 2007, "When individual behaviour matters: homogeneous and network models in epidemiology." *J. R. Soc. Interface*, 4: pp. 879-891.

[3] M. Barthelemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, 2005, "Dynamical patterns of epidemic outbreaks in complex heterogeneous networks." *Journal of Theoretical Biology*, 235: pp. 275-288.

[4] E. Beretta, T. Hara, W. Ma, and Y. Takeuchi, 2001, "Global asymptotic stability of an SIR epidemic model with distributed time delay." *Nonlinear Anal.*, 47: pp. 4107-4115.

[5] V. Colizza, A. Barrat, M. Barthelemy, and A. Vespignani, 2006, "The role of the airline transportation network in the prediction and predictability of global epidemics." *PNAS*, 103: pp. 2015-2020.

[6] H. W. Hethcote, 1976, " Qualitative analyses of communicable disease models." *Math. Biosci.*, 7: pp. 335-356.

[7] T. House, G. Davies, L. Danon, and M. J. Keeling, 2009, "A motif-based approach to network epidemics." *Bulletin of Mathematical Biology*, 71: pp. 1693-1706.

[8] E. M. Jin, M. Girvan, and M. E. J. Newman, 2001, "Structure of growing social networks." *Physical Review E*, 64, 046132.

[9] L. Huang, K. Park, and Y.-C. Lai, 2006, "Information propagation on modular networks." *Physical Review E*, 73, 035103 (R).

[10] M. J. Keeling and K. T. D. Eames, 2005, "Networks and epidemic models." *J. R. Soc. Interface*, 2: pp. 295-307.

[11] C.-H. Li, 2009, *Analysis of synchronization in nonlinearly coupled dynamical networks*, PhD Dissertation, Department of Mathematics, National Central University, Taiwan.

[12] C.-H. Li, C.-C. Tsai, and S.-Y. Yang, 2011, "Permanence of an SIR epidemic model with density dependent birth rate and distributed time delays." *Appl. Math. Comput.*, 218: pp. 1682-1693.

[13] C.-H. Li, C.-C. Tsai, and S.-Y. Yang, 2012, "Analysis of the permanence of an SIR epidemic model with logistic process and distributed time delay." *Commun. Nonlinear Sci. Numer. Simulat.*, 17: pp. 3696-3707.

[14] M. Y. Li, Z. Shuai, and C. Wang, 2010, “Global stability of multi-group epidemic models with distributed delays.” *J. Math. Anal. Appl.*, 361: pp. 38-47.

[15] Z. Liu and B. Hu, “Epidemic spreading in community networks.” *Europhys. Lett.*, 72 (2005) : pp. 315-321.

[16] W. Ma, M. Song, and Y. Takeuchi, 2004, “Global stability of an SIR epidemic model with time delay.” *Appl. Math. Letter.*, 17, pp. 1141-1145..

[17] W. Ma, Y. Takeuchi, T. Hara, and E. Beretta, 2002, “ Permanence of an SIR epidemic model with distributed time delays.” *Tohoku Math. J.*, 54, pp. 581-591.

[18] L. A. Meyers, B. Pourbohloul, M. E. J. Newman, D. M. Skowronski, and R. C. Brunham, 2005, “Network theory and SARS: predicting outbreak diversity.” *Journal of Theoretical Biology*, 232: pp. 71-81.

[19] Y. Moreno, R. Pastor-Satorras, and A. Vespignani1, 2002, “ Epidemic outbreaks in complex heterogeneous networks.” *Eur. Phys. J. B*, 26: pp. 521-529.

[20] M. E. J. Newman, 2010, *Networks: An Introduction*, Oxford University Press Inc., New York.

[21] R. Pastor-Satorras and A. Vespignani, 2001, “Epidemic spreading in scale-free networks.” *Physical Review Letters*, 86: pp. 3200-3203.

[22] R. Pastor-Satorras and A. Vespignani, 2002, “ Epidemic dynamics in finite size scale-free networks.” *Phys. Rev. E*, 65, 035108 (R).

[23] L. Wang and G.-Z. Dai, 2008, “Global stability of virus spreading in complex heterogeneous networks.” *SIAM J. Appl. Math.*, 68: pp. 1495-1502.

[24] X.-J. Xu, H.-O. Peng, X.-M. Wang, and Y.-H. Wang, 2006, “Epidemic spreading with time delay in complex networks.” *Physica A*, 367: pp. 525-530.

[25] F. Zhang and X.-Q. Zhao, 2007, “A periodic epidemic model in a patchy environment.” *J. Math. Anal. Appl.*, 325: pp. 496-516.

[26] G. Zhu, X. Fu, and G. Chen, 2012, “Global attractivity of a network-based epidemic SIS model with nonlinear infectivity.” C*ommun. Nonlinear Sci. Numer. Simulat.*, 17: pp. 2588-2594.