

Introduction to Data Science: Homework 3

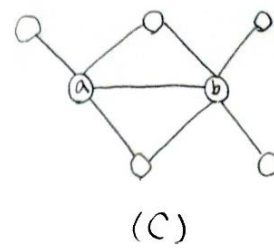
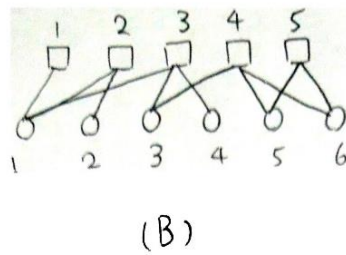
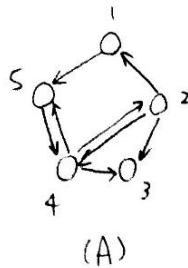
醫工系 107 吳旻昇 F94036089

1. Give **one** real-life example of each of the following types of networks; then, briefly describe **one** empirical technique that could be used to measure the structure of each of the following networks (i.e., to fully determine the positions of all the edges), and **one** phenomenon regarding the network that you are interested in which network analysis methods may help to understand:
 - (a) An acyclic (or approximately acyclic) directed network
 - (1) 家族譜系圖是有方向性的非循環網路結構。
 - (2) 透過口頭調查、歷史考證、基因分析等方式可以獲得家族譜系圖。
 - (3) 分析家族譜系圖可以知道「彼此之間的血緣關係」。
 - (b) A cyclic directed network
 - (1) 食物鏈是有方向性的循環網路結構
 - (2) 透過生態觀察、記錄、採集等方式可以獲得特定生態系中的食物鏈。
 - (3) 分析食物鏈對於「那些物種對於此生態系的平衡非常重要？」的環境保育議題很有幫助。
 - (c) A tree (or approximate tree)
 - (1) 物種演化樹是一種樹狀結構。
 - (2) 透過化石與現存物種的性狀考證、元素定年、DNA 序列比對等方法可以建立物種演化樹。
 - (3) 分析物種演化樹不僅可以知道時間上的「某物種的演化史以及其與相近物種的親緣關係」，更能由其空間上的分布探討「該物種的活動或遷移方式」。
 - (d) A planar (or approximately planar) network
 - (1) 平面道路路網是可以畫在平面上並且讓不同邊互不交疊的網路結構。
 - (2) 透過衛星遙測、地理資訊系統 (GIS) 分析等可以獲得平面道路路網。
 - (3) 分析平面道路路網可以釐清「新道路規劃、交通號誌安排、施工影響」等問題。
 - (e) A bipartite network
 - (1) NHK 紅白歌合戰是一種 bipartite network 。
 - (2) 透過蒐集 NHK 電視台的資料、收視率調查可以獲得紅白歌合戰的 bipartite network 。
 - (3) 分析紅白歌合戰 bipartite network 可以知道「哪一組藝人對於自己隊的收視率貢獻度最大？」。

(f) A temporal network (or sequence of edges)

- (1) 流行性感冒的傳播是一種 temporal network 。
- (2) 透過蒐集病歷資料、傳染病通報系統可以建立流行性感冒傳播的 temporal network 。
- (3) 分析流行性感冒傳播的 temporal network 可以知道「要如何有效防治傳染病的傳播？」

2. Consider the following three networks:



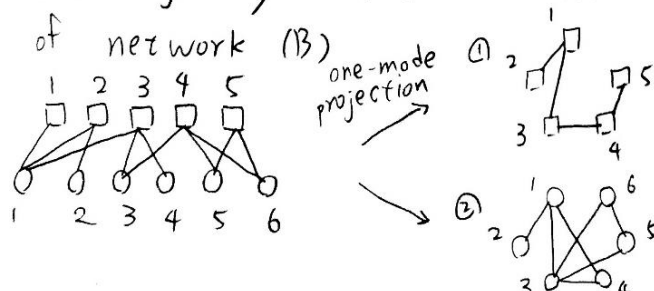
(a) Give the adjacency matrix for network (A)

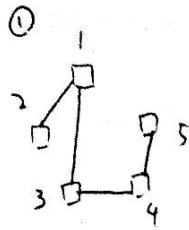
A	1	2	3	4	5
1	0	0	0	0	1
2	1	0	1	1	0
3	0	0	0	0	0
4	0	1	1	0	1
5	0	0	0	1	0

(b) Give adjacency list for network (A)

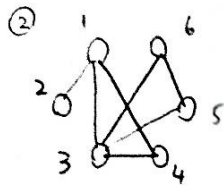
A	
1	$\rightarrow \{(5, 1)\}$
2	$\rightarrow \{(1, 1), (3, 1), (4, 1)\}$
3	$\rightarrow \{\}$
4	$\rightarrow \{(2, 1), (3, 1), (5, 1)\}$
5	$\rightarrow \{(4, 1)\}$

(c) Give adjacency matrices for both one-mode projections





0	1	2	3	4	5
1	0	1	1	0	0
2	1	0	0	0	0
3	1	0	0	1	0
4	0	0	1	0	1
5	0	0	0	1	0



0	1	2	3	4	5	6
1	0	1	1	1	0	0
2	1	0	0	0	0	0
3	1	0	0	1	1	1
4	1	0	1	0	0	0
5	0	0	1	0	0	1
6	0	0	1	0	1	0

(d) What is the cosine similarity of vertices a and b in network (c)?

Cosine Similarity: $\text{Cosine}(v_i, v_j) = \frac{|N_i \cap N_j|}{\sqrt{|N_i| \cdot |N_j|}}$

v_i, v_j : 2 個 不同 的 vertices

N : 某 vertex 的 neighbors

$N_i \cap N_j$: v_i 和 v_j 的 common neighbors

$\text{cosine}(a, b) = \frac{2}{\sqrt{4 \cdot 5}} = 0.4472$

Collaboration: No one, only by myself.