POLITECNICO
MILANO 1863

COMPLESSITA' NEI SISTEMI E NELLE RETI

Prof. C. Piccardi
Appello del 31/1/2022

COGNOME:________________________ NOME: _______________________

Codice Persona: _____________________ Corso di laurea (INF, MTM, ...): ______

Firma dello studente: ___________________________ Visto del docente:______

5	5	5	5	5	5

Voto totale
30

AVVERTENZE

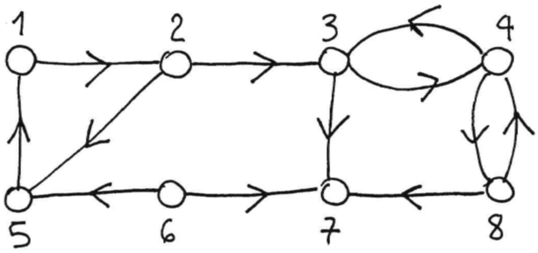
- Non è consentito consultare alcun tipo di materiale (libri, appunti, smartphone, ecc.).

- Oltre alla pertinenza e completezza della risposta, sono valutati anche **ordine, chiarezza e rigore formale**.
- Le soluzioni devono essere riportate solo sui fogli allegati.
- Lo studente è tenuto a prendere visione delle **modalità d'esame dettagliate** alla pagina web del corso (http://piccardi.faculty.polimi.it/csr.html).

Rispondere con precisione e rigore formale, riportando le eventuali formule utilizzate e svolgendo i calcoli richiesti.

Problema 1 (5 punti)

Si consideri la rete diretta, non pesata, rappresentata in figura.



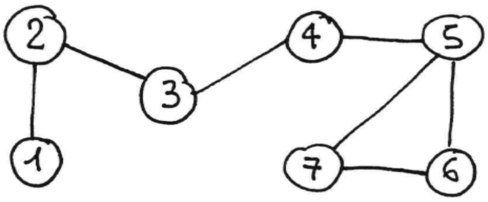
a) Determinare (se esistono) le componenti SCC, IN e OUT.
b) Determinare la matrice delle probabilità di transizione di un random walker, aggiungendo autoanelli alla rete solo quando strettamente necessario.
c) Sulla rete definita al punto b), determinare la random walk centrality di ciascun nodo (*è possibile rispondere su base intuitiva, senza svolgere calcoli*).

Soluzione [se necessario proseguire sul retro]:

Rispondere con precisione e rigore formale, riportando le eventuali formule utilizzate e svolgendo i calcoli richiesti.

Problema 2 (5 punti)

Si consideri la rete non diretta, non pesata, rappresentata in figura.



Si considerino le tre partizioni seguenti:

$$P_1 = \{\{1,2\},\{3,4\},\{5,6,7\}\}$$
$$P_2 = \{\{1,2,3,4\},\{5,6,7\}\}$$
$$P_3 = \{\{1,2,3\},\{4,5,6,7\}\}$$

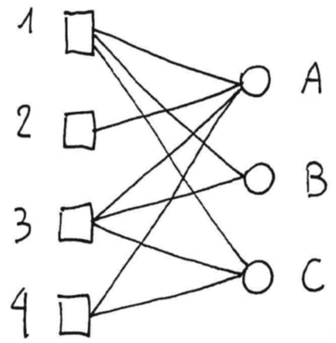
a) Determinare la migliore partizione secondo il criterio della massima modularità.
b) Determinare la migliore partizione secondo il criterio della probabilità di persistenza (*selezionare la partizione con la probabilità di persistenza media più elevata*).

Soluzione [se necessario proseguire sul retro]:

Rispondere con precisione e rigore formale, riportando le eventuali formule utilizzate e svolgendo i calcoli richiesti.

Problema 3 (5 punti)

Si consideri la rete bipartita in figura, formata da 4 utenti (1,2,3,4) e 3 oggetti (A,B,C).



a) Scrivere la matrice di incidenza della rete.
b) Determinare le matrici dei pesi (o matrici di adiacenza pesate) relative, rispettivamente, alle proiezioni sull'insieme degli utenti e sull'insieme degli oggetti.
c) Nell'ottica di realizzare un recommender system basato su *collaborative filtering*, si determini a quale degli utenti che non possiedono l'oggetto B è più opportuno raccomandarlo.

Soluzione [se necessario proseguire sul retro]:

Rispondere con precisione e rigore formale, specificando il significato di tutte le variabili introdotte ed elencando le eventuali ipotesi necessarie. Non sono richiesti commenti, motivazioni, dimostrazioni, ecc..

Domanda 4 (5 punti)

Stochastic block-model: definizione e casi notevoli.

<u>Rispondere con precisione e rigore formale, specificando il significato di tutte le variabili introdotte ed elencando le eventuali ipotesi necessarie. Non sono richiesti commenti, motivazioni, dimostrazioni, ecc..</u>

Domanda 5 (5 punti)

Immunizzazione in un processo SIS su reti omogenee ed eterogenee: definizione e proprietà.

<u>Indicare la risposta esatta (che è sempre unica): non è richiesta giustificazione (risposta esatta = +1, risposta errata = − 0.5, risposta non data = 0)</u>

In una rete non diretta, non pesata, gli autovalori della matrice Laplaciana

[1] possono essere reali o complessi, con parte reale non negativa
[2] sono tutti reali positivi
[3] sono tutti reali non negativi
[4] sono tutti reali negativi o nulli

In una rete diretta, la random walk centrality è sicuramente ben definita (cioè esiste unica per ogni nodo)

[1] se la rete è completa
[2] se la rete non ha nodi isolati
[3] se la rete ha al più due componenti
[4] se la rete è connessa in senso debole

In una rete non diretta, pesata, la random walk centrality di un nodo

[1] coincide con il suo grado
[2] è proporzionale al suo grado
[3] coincide con la sua strength
[4] è proporzionale alla sua strength

In un processo SIS su rete eterogenea, nella quale $\langle k^2 \rangle \to \infty$ per $N \to \infty$, la soglia epidemica

[1] è finita e vale $\gamma/\langle k \rangle$
[2] tende a ∞ per $N \to \infty$
[3] tende a 0 per $N \to \infty$
[4] non è possibile affermare nulla senza conoscere la specifica topologia della rete

Nel corso del laboratorio Gephi si è analizzata la seguente rete:

[1] coautori di articoli di area ecologia
[2] coautori di articoli di area neuroscienze
[3] mobilità per via aerea tra i cittadini delle nazioni europee
[4] mobilità tra gli abitanti dei comuni italiani

1. a) $SCC_1 = \{1, 2, 5\}$

$SCC_2 = \{3, 4, 8\}$

$IN = \{6\}$

$OUT = \{7\}$ (il link $6 \to 7$ è un "tube")

b) $$p_{ij} = \frac{a_{ij}}{k_i^{out}}$$

$$P = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \end{vmatrix}$$

autoanello poiché $k_7^{out} = 0$

c) Si deve risolvere il sistema di equazioni $\pi = \pi P$.

Si osserva che:

- da ogni nodo esiste un percorso verso 7, che quindi verrà raggiunto con probabilità 1
- da 7 non esistono percorsi in uscita, ma solo l'autoanello

Per $t \to \infty$, quindi, il random walker si troverà nel nodo 7, per cui:

$$\pi = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

② a)

$$A = \begin{array}{c|ccccccc|} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 5 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 6 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array} \qquad [k_i k_j] = \begin{vmatrix} 1 & 2 & 2 & 2 & 3 & 2 & 2 \\ 2 & 4 & 4 & 4 & 6 & 4 & 4 \\ 2 & 4 & 4 & 4 & 6 & 4 & 4 \\ 2 & 4 & 4 & 4 & 6 & 4 & 4 \\ 3 & 6 & 6 & 6 & 9 & 6 & 6 \\ 2 & 4 & 4 & 4 & 6 & 4 & 4 \\ 2 & 4 & 4 & 4 & 6 & 4 & 4 \end{vmatrix}$$

$$[k_i] = \begin{vmatrix} 1 & 2 & 2 & 2 & 3 & 2 & 2 \end{vmatrix}$$

$$2L = \sum_i k_i = 14 \qquad Q = \frac{1}{2L} \sum_{c \in P} \sum_{i,j \in c} \left(a_{ij} - \frac{k_i k_j}{2L} \right)$$

$$Q_1 = \frac{1}{14} \left\{ \left[2 - \frac{9}{14} \right] + \left[2 - \frac{16}{14} \right] + \left[6 - \frac{49}{14} \right] \right\} = 0.3367$$

$$Q_2 = \frac{1}{14} \left\{ \left[6 - \frac{49}{14} \right] + \left[6 - \frac{49}{14} \right] \right\} = 0.3571$$

P_2 è la partizione a massima modularità.

$$Q_3 = \frac{1}{14} \left\{ \left[4 - \frac{25}{14} \right] + \left[8 - \frac{81}{14} \right] \right\} = 0.3163$$

b) $$\alpha_c = \frac{\sum_{i \in c} k_i^{INT}}{\sum_{i \in c} k_i}$$

$$P_1: \quad \alpha_1 = \frac{1+1}{1+2} = \frac{2}{3}, \quad \alpha_2 = \underbrace{\frac{1+1}{2+2}}_{=1/2}, \quad \alpha_3 = \frac{2+2+2}{3+2+2} = \frac{6}{7} \quad : \bar{\alpha} = 0.6746$$

$$P_2: \quad \alpha_1 = \frac{1+2+2+1}{1+2+2+2} = \frac{6}{7}, \quad \alpha_2 = \frac{6}{7} \quad : \quad \bar{\alpha} = 0.8571$$

P_2 è la part. con la max prob. di pers. media

$$P_3: \quad \alpha_1 = \frac{1+2+1}{1+2+2} = \frac{4}{5}, \quad \alpha_2 = \frac{1+3+2+2}{2+3+2+2} = \frac{8}{9} \quad : \bar{\alpha} = 0.8444$$

③ a)

$$B = \begin{array}{c|ccc} & A & B & C \\ \hline 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ 4 & 1 & 0 & 1 \end{array}$$

b) $$W' = BB^T - diag(BB^T) = \begin{vmatrix} 3 & 1 & 3 & 2 \\ 1 & 1 & 1 & 1 \\ 3 & 1 & 3 & 2 \\ 2 & 1 & 2 & 2 \end{vmatrix} - \begin{vmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{vmatrix} =$$

$$= \begin{vmatrix} 0 & 1 & 3 & 2 \\ 1 & 0 & 1 & 1 \\ 3 & 1 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{vmatrix}$$

$$W'' = B^TB - diag(B^TB) = \begin{vmatrix} 4 & 2 & 3 \\ 2 & 2 & 2 \\ 3 & 2 & 3 \end{vmatrix} - \begin{vmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 3 \\ 2 & 0 & 2 \\ 3 & 2 & 0 \end{vmatrix}$$

c) $$\hat{b}_{ij} = \frac{\sum_{e=1}^{m} s_{ie} b_{ej}}{\sum_{e=1}^{m} s_{ie}}$$

Le similarità s_{ie} coincidono di fatto con gli elementi di W'!

$$\hat{b}_{2B} = \frac{s_{21} b_{1B} + s_{23} b_{3B} + s_{24} b_{4B}}{s_{21} + s_{23} + s_{24}} = \frac{1 \cdot 1 + 1 \cdot 1 + 1 \cdot 0}{1 + 1 + 1} = \frac{2}{3}$$

$$\hat{b}_{4B} = \frac{s_{41} b_{1B} + s_{42} b_{2B} + s_{43} b_{3B}}{s_{41} + s_{42} + s_{43}} = \frac{2 \cdot 1 + 1 \cdot 0 + 2 \cdot 1}{2 + 1 + 2} = \frac{4}{5}$$

$\hat{b}_{4B} > \hat{b}_{2B}$, per cui l'oggetto B è raccomandato prioritariamente all'utente 4.