

TD - Network Science

Mars 2024 – Riadh DHAOU

Compter : **1 heure 30 de travail**- Nombre de pages : 5 pages

Consigne :

Compléter les réponses aux questions de ce TD et rendre pour le **rendre le vendredi 21 Mars 2024** à 19h00.

Une correction sera disponible le même jour.

Nom :SORRES / HOUOT.....Prénom : Antonin / Léa

2SN-R

Questions de cours (4 points)

Q1) Expliquer brièvement le principe d'attachement préférentiel dans un réseau complexe évolutif.

Dans un réseau complexe évolutif, le principe d'attachement préférentiel désigne le fait que quand un nouveau nœud est créé il est préférentiellement attaché à un autre nœud fortement connecté. Par exemple, la création d'un nouvel aéroport qui aura une ligne vers un aéroport plus gros.

Q2) Comment la robustesse d'un réseau est-elle mesurée ? Proposez quelques métriques qui permettent de mesurer la robustesse.

La robustesse est mesurée de manière empirique, en effet, elle est mesurée en regardant la connectivité globale du réseau lorsqu'on y enlève des éléments (lien, nœuds). Comme métriques il y a :

- coefficient de clustering ;
- distance moyenne entre deux nœuds.

Q3) Que reflète la métrique modularité ? Comment cette métrique est-elle utilisée ?

La métrique modularité est le score reflétant la capacité qu'a un réseau à se sous-diviser en communauté.

Cette métrique est utilisé pour comparer des méthodes de détection de communauté.

Q4) Quel(s) modèle(s) peut-on utiliser pour étudier la dissémination d'une rumeur dans un réseau social ?

Il existe 3 grands modèles : SI (Susceptible Infected), SIS (Susceptible Infected Susceptible) et SIR (Susceptible Infected Recovered).

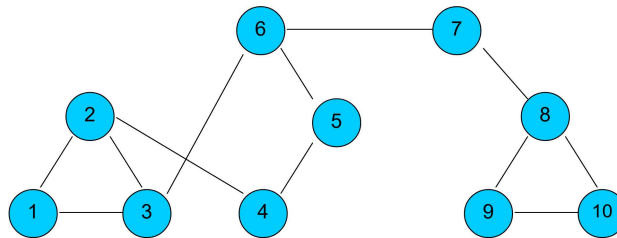
Le modèle SI : il y a une rumeur qui se propage et tout le monde peut commencer à la croire et il change jamais d'avis.

Le modèle SIS : on peut remettre en question la rumeur et ne plus la croire mais d'autres arguments peuvent faire qu'on la croit à nouveau.

Le modèle SIR : on ne croit plus la rumeur et rien ne nous fera changer d'avis.

Exercice: Communautés dans un réseau social

Q1) Donner le nombre chromatique du graphe suivant:



Le nombre chromatique du graphe est 3.

Le graphe précédent modélise un réseau social composé de 10 individus. Nous voulons dans la suite déterminer les degrés de centralité des individus dans ce réseau. Trois définitions de la centralité sont données.

Centralité d'intermédiation ($C_{\text{intermédiation}}$): La centralité d'intermédiation compte le nombre de fois où un nœud agit comme un point de passage le long du plus court chemin entre deux autres nœuds. Elle a été présentée comme une mesure pour quantifier le contrôle d'un humain sur la communication entre d'autres humains dans un réseau social par Linton Freeman. Dans sa conception, les sommets qui ont une forte probabilité d'apparaître sur un court chemin choisi au hasard entre deux sommets choisis au hasard ont une haute intermédiation.

L'intermédiation d'un sommet v dans $G = (V, E)$ avec $|V|$ sommets est calculée comme suit:

- Pour chaque couple (s, t) , on calcule les plus courts chemins les reliant.
- Pour chaque couple (s, t) , on détermine la proportion de plus courts chemins qui passent par le sommet en question, ici v .
- On somme cette fraction sur tous les couples (s, t) de sommet.

De façon plus concise, l'intermédiation peut être représentée par :

$$C(v) = \sum_{s \neq v, t \neq v} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

où $\sigma_{s,t}$ est le nombre total de plus courts chemins du sommet s au sommet t et $\sigma_{s,t}(v)$ est le nombre de tels chemins qui passent par v .

Notons que, dans le graphe étudié, il existe un seul plus court chemin entre chaque couple de sommets pris au hasard.

Centralité de proximité ($C_{\text{proximité}}$): Dans un graphe connexe il y a une mesure de distance naturelle entre paires de nœuds, définie par la longueur de leurs plus courts chemins. L'excentricité d'un nœud x est définie comme la somme des distances à tous les autres nœuds, et la proximité est définie par Bavelas comme l'inverse de l'éloignement.

$$C(x) = \frac{1}{\sum_y d(y, x)}$$

Centralité de prestige (C_{Prestige}): Cette métrique reflète l'importance d'un nœud dans un réseau social. Cette mesure assigne des scores relatifs à chacun des nœuds du réseau en se basant sur le principe que les connexions vers les nœuds ayant les scores les plus élevés, contribuent davantage au score du nœud en question que des connexions égales mais à de plus bas score. En d'autres termes, c'est la somme des degrés des voisins à deux sauts du nœud (directs et indirects).

La matrice suivante donne la longueur de l'ensemble des plus courts chemins de ce graphe.

	1	2	3	4	5	6	7	8	9	10
1	0	1	1	2	3	2	3	4	5	5
2	1	0	1	1	2	2	3	4	5	5
3	1	1	0	2	2	1	2	3	4	4
4	2	1	2	0	1	2	3	4	5	5
5	3	2	2	1	0	1	2	3	4	4
6	2	2	1	2	1	0	1	2	3	3
7	3	3	2	3	2	1	0	1	2	2
8	4	4	3	4	3	2	1	0	1	1
9	5	5	4	5	4	3	2	1	0	1
10	5	5	4	5	4	3	2	1	1	0

La matrice suivante présente tous les plus courts chemins

	1	2	3	4	5	6	7	8	9	10
1	1	1.2	1.3	1.2.4	1.2.4.5	1.3.6	1.3.6.7	1.3.6.7.8	1.3.6.7.8.9	1.3.6.7.8.10
2		2	2.3	2.4	2.4.5	2.3.6	2.3.6.7	2.3.6.7.8	2.3.6.7.8.9	2.3.6.7.8.10
3			3	3.2.4	3.6.5	3.6	3.6.7	3.6.7.8	3.6.7.8.9	3.6.7.8.10
4				4	4.5	4.5.6	4.5.6.7	4.5.6.7.8	4.5.6.7.8.9	4.5.6.7.8.10
5					5	5.6	5.6.7	5.6.7.8	5.6.7.8.9	5.6.7.8.10
6						6	6.7	6.7.8	6.7.8.9	6.7.8.10
7							7	7.8	7.8.9	7.8.10
8								8	8.9	8.10
9									9	9.10
10										10

Les nombres de plus courts chemins, passant par chaque nœud, sont donnés comme suit :

	1	2	3	4	5	6	7	8	9	10
Nombre	0	3	10	2	5	21	18	14	0	0

Q2) Compléter le tableau suivant et déduire les nœuds centraux du réseau:

	1	2	3	4	5	6	7	8	9	10
$C_{\text{Intermédianité}}$	0	3/45	10/45	2/45	5/45	21/45	18/45	14/45	0	0
$C_{\text{Proximité}}$	1/26	1/24	1/20	1/25	1/22	1/17	1/19	1/23	1/30	1/30
C_{Prestige}	11	12	15	13	13	17	15	9	5	5

Les nombres de plus courts chemins, passant par chaque lien, sont donnés comme suit :

	1-2	1-3	2-3	2-4	3-6	4-5	5-6	6-7	7-8	8-9	8-10	9-10
Nombre chemins	3	6	6	5	16	8	10	24	21	8	8	1

Q3) Divisez l'ensemble du réseau en 2, 3 et 4 sous-communautés, respectivement.

AIDE : utiliser une méthode qui divise la communauté en plusieurs sous-communautés en supprimant en priorité les liens reliant différentes sous-communautés (les liens parcourus par le plus grand nombre de plus-courts chemins).

2 communautés : 1, 2, 3, 4, 5, 6 + 7, 8, 9, 10

3 communautés : 1, 2, 3 + 4, 5, 6 + 7, 8, 9, 10

4 communautés : 1, 2, 3 + 4, 5 + 6, 7 + 8, 9, 10

Q4) Calculer la modularité pour chaque partition proposée. Quel est le meilleur partitionnement ?

RAPPEL : La modularité d'une partition c de la communauté globale en n_c sous-communautés est :

$$M_c = \sum_{c=1}^{n_c} \left[\frac{L_c}{L} - \left(\frac{k_c}{2L} \right)^2 \right]$$

où L_c est le nombre de liens au sein de la sous-communauté c , L est le nombre de liens au sein de la communauté globale et k_c est la somme des degrés des nœuds de la sous-communauté c .

2 communautés :

1ère communauté {1, 2, 3, 4, 5, 6}

$$L_c = 7$$

$$k_c = 15$$

2ème communauté {7, 8, 9, 10}

$$L_c = 4$$

$$k_c = 9$$

$$M_c = 7/12 - (15/24)^2 + 4/12 - (9/24)^2 = 0.385$$

3 communautés :

1ère communauté {1, 2, 3}

$$L_c = 3$$

$$k_c = 8$$

2ème communauté {4, 5, 6}

$$L_c = 2$$

$$k_c = 7$$

3ème communauté {7, 8, 9, 10}

$$L_c = 4$$

$$k_c = 9$$

$$M_c = 3/12 - (8/24)^2 + 2/12 - (7/24)^2 + 4/12 - (9/24)^2 = 0.413$$

4 communautés :

1ère communauté {1, 2, 3}

$$L_c = 3$$

$$k_c = 8$$

2ème communauté {4, 5}

$$L_c = 1$$

$$k_c = 4$$

3ème communauté {6, 7}

$$L_c = 1$$

$$k_c = 5$$

4ème communauté {8, 9, 10}

$$L_c = 3$$

$$k_c = 7$$

$$M_c = 3/12 - (8/24)^2 + 1/12 - (4/24)^2 + 1/12 - (5/24)^2 + 3/12 - (7/24)^2 = 0.399$$

Le meilleur partitionnement est celui en 3 communauté, en effet il a une métrique de modularité plus grande que celle en 2 communautés et celle en 4 communautés.