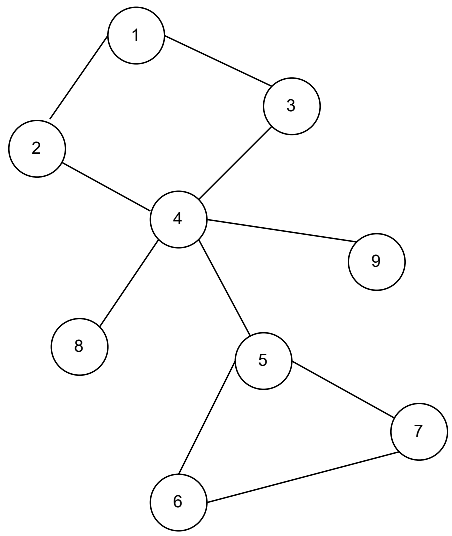
Social Network Analysis Exam - Solutions

Unknown EDEMM author

17-02-2023

Θέμα 1



Degree centrality:

$$C_D(1) = C_D(2) = C_D(3) = C_D(6) = C_D(7) = 2$$
$$C_D(4) = 5$$
$$C_D(8) = C_D(9) = 1$$
$$C_D(5) = 3$$

Closeness centrality:

$$C_P(i) = \frac{1}{\sum_j d(i,j)}$$
$$C_P(1) = \frac{1}{1+1+2+3+3+3+4+4} = \frac{1}{21}$$
$$C_P(2) = \frac{1}{1+2+1+2+2+2+3+3} = \frac{1}{16} = C_P(3) \text{ λόγω συμμετρίας}$$
$$C_P(4) = \frac{1}{2+1+1+1+1+1+2+2} = \frac{1}{11}$$
$$C_P(8) = \frac{1}{3+2+2+1+2+2+3+3} = \frac{1}{18} = C_P(9) \text{ λόγω συμμετρίας}$$
$$C_P(5) = \frac{1}{3+2+2+1+2+2+1+1} = \frac{1}{14}$$
$$C_P(6) = \frac{1}{4+3+3+2+3+3+1+1} = \frac{1}{20} = C_P(7) \text{ λόγω συμμετρίας}$$

Ο κόμβος με το μεγαλύτερο closeness centrality είναι ο 4 όπως και θα αναμέναμε, καθώς είναι ο πιο κεντρικός χωρικά κόμβος.

Betweenness centrality:

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

$$C_B(1) = \underbrace{\frac{1}{2}}_{2 \to 3}$$

$$C_B(2) = \underbrace{6 \cdot \frac{1}{2}}_{1 \to 4,5,6,7,8,9} = 3 = C_B(3) \text{ λόγω συμμετρίας}$$

$$C_B(4) = \underbrace{3 \cdot 5}_{1,2,3 \to 5,6,7,8,9} + \underbrace{3 \cdot 2}_{5,6,7 \to 8,9} = 15 + 6 = 21$$

$$C_B(5) = \underbrace{2 \cdot 6}_{6,7 \to 1,2,3,4,8,9} = 12$$

$$C_B(5) = 0 = C_B(7) \text{ λόγω συμμετρίας}$$

$$C_B(8) = 0 = C_B(9) \text{ λόγω συμμετρίας}$$

Θέμα 2

(α)

Τοπολογία	**Σχεσιακή/χωρική**	**Τύπος τοπολογίας**	**Ντετερμινιστική/στοχαστική**
1	σχεσιακή	Regular graph	ντετερμινιστική
2	σχεσιακή	Scale free	στοχαστική
3	χωρική	Random geometric graph	στοχαστική
4	σχεσιακή	Random graph	στοχαστική

(β)

Τύπος Δικτύου	**Παράδειγμα Πραγματικού Δικτύου**
Κανονικό	Κρυσταλλικό πλέγμα
Small-world	Peer to peer networks (eg torrent file sharing)
Τυχαίο (ER)	Random user networks (WhatsApp)
Scale-free	Social Influence Networks (Twitter)
Τυχαιο Γεωμετρικό	Geographical Information System

Θέμα 3

Το πρόβλημα αυτό συναντάται στη βιβλιογραφία ως το Gambler's ruin problem.

$$
\begin{aligned}
&p(N) = 1\\
&p(0) = 0\\
&p(x) = \frac{1}{2}p(x-1) + \frac{1}{2}p(x+1)\ ,\ 0 < x < N, \text{το οποίο ξαναγράφεται}\\
&p(x+1) - p(x) = p(x) - p(x-1)
\end{aligned}
$$

Αν διαδοχικά στην προηγούμενη σχέση αντικαταστήσουμε το x με $x-1$, $x-2$, κ.ο.κ. έχουμε:

$$
\begin{aligned}
&p(x) - p(x-1) = p(x-1) - p(x-2)\\
&p(x-1) - p(x-2) = p(x-2) - p(x-3)\\
&...\\
&p(3) - p(2) = p(2) - p(1)\\
&p(2) - p(1) = p(1) - p(0)
\end{aligned}
$$

όπου βλέπουμε πως όλες αυτές οι διαφορές είναι ίσες. Αν θέσουμε $m = p(x) - p(x-1)$, τότε λοιπόν:

$$
\begin{aligned}
&m = p(x) - p(x-1)\\
&m = p(x-1) - p(x-2)\\
&m = p(x-2) - p(x-3)\\
&...\\
&m = p(3) - p(2)\\
&m = p(2) - p(1)\\
&m = p(1) - p(0)
\end{aligned}
$$

οπότε αθρόιζοντας κατά μέλη παίρνουμε:

$$m \cdot x = p(x) - p(0), \text{ ή ισοδύναμα:}$$

$$p(x) = p(0) + m \cdot x$$

Όμως $p(0) = 0$ και $p(N) = 1$ άρα για $x = N$ έχουμε:

$$p(N) = p(0) + m \cdot N$$

$$1 = m \cdot N$$

$$m = \frac{1}{N}$$

Οπότε τελικά:

$$p(x) = \frac{x}{N}\ ,\ x \in 0, 1, ..., N$$

Τέλος για $N = 5$, τα αποτελέσματα για τα διάφορα $p(x)$ είναι προφανή.

Θέμα 4

1.

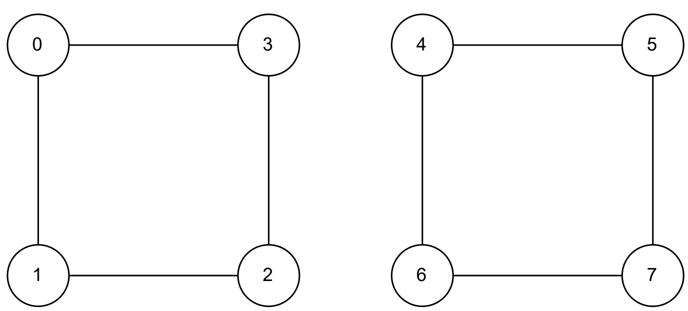
$$C_B(0,3) = \underbrace{\frac{1}{2}}_{0\to 2} + \underbrace{1}_{0\to 3} + \underbrace{\frac{2}{3}}_{0\to 4} + \underbrace{\frac{3}{4}}_{0\to 5} + \underbrace{1}_{0\to 6} + \underbrace{1}_{0\to 7} + \underbrace{\frac{1}{2}}_{1\to 3} + \underbrace{\frac{1}{3}}_{1\to 6} + \underbrace{\frac{1}{4}}_{1\to 7} = 6$$

$$C_B(3,6) = \underbrace{\frac{1}{3}}_{0\to 4} + \underbrace{\frac{2}{4}}_{0\to 5} + \underbrace{1}_{0\to 6} + \underbrace{1}_{0\to 7} + \underbrace{\frac{2}{3}}_{1\to 6} + \underbrace{\frac{2}{4}}_{1\to 7} + \underbrace{\frac{1}{2}}_{2\to 6} + \underbrace{\frac{1}{3}}_{2\to 7} + \underbrace{\frac{1}{2}}_{3\to 4} + \underbrace{\frac{2}{3}}_{3\to 5} + \underbrace{1}_{3\to 6} + \underbrace{1}_{3\to 7} = 8$$

Λόγω συμμετρίας θα είναι $C_B(3,6) = C_B(2,4) = 8$, το οποίο θα είναι το μέγιστο edge betweenness centrality στο γράφο. Ένας εναλλακτικός γρήγορος τρόπος υπολογισμού είναι πως δεξιά και αριστερά βλέπουμε δύο communities στο γράφο από 4 κόμβους η καθεμία, οπότε αν είχαμε μία ακμή-γέφυρα θα είχαμε $4 \cdot 4 = 16$ κεντρικότητα, όμως μια και έχουμε δύο ακμες-γέφυρες, αυτή σπάει σε 8 και 8 ανάμεσα στις δύο αυτές συμμετρικές γέφυρες.

2.

Ο Girvan-Newman βρίσκει τις ακμές με τη μέγιστη ενδιαμεσική κεντρικότητα και σε κάθε βήμα του αλγορίθμου αφαιρεί εκείνη με την υψηλότερη κεντρικότητα. Ύστερα επανυπολογίζονται οι κεντρικότητες στο γράφο που έχει προκύψει και επαναλαμβάνεται η διαδικασία έως ότου διαχωριστεί σε κοινότητες ο γράφος. Όποια εκ των δύο ακμών (3,6) και (2,4) αν αφαιρέσουμε στην 1η επανάληψη, στη 2η θα αφαιρεθεί η δεύτερη εξ αυτών έχοντας την εκ νέου υπολογισμένη υψηλότερη κεντρικότητα (όντας γέφυρα). Οι δύο κοινότητες που προκύπτουν είναι τελικά:



Κοινότητα 1 (C1) Κοινότητα 2 (C2)

3.
Για την παραπάνω διαμέριση έχουμε $m = 8$ συνολικά ακμές.

$$Q = \frac{1}{2m} \sum_{l=1}^{k} \sum_{i,j \in C_l} \left(A_{ij} - \frac{d_i d_j}{2m} \right)$$
$$= \frac{1}{16} \left(\underbrace{\sum_{i,j \in C_1} \left(A_{ij} - \frac{d_i d_j}{16} \right)}_{C} + \underbrace{\sum_{i,j \in C_2} \left(A_{ij} - \frac{d_i d_j}{16} \right)}_{D} \right)$$

Λόγω συμμετρίας οι ποσότητες $C = D$, ενώ για όλους τους κόμβους ισχύει $d_i = 2$, $i \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ οπότε $\frac{d_i d_j}{16} = \frac{1}{4}$. Για τον υπολογισμό του C που αφορά την πρώτη κοινότητα C_1 παίρνουμε όλα τα ζεύγη κόμβων που είναι 12 εν προκειμένω δηλαδή τα $(0,1), (0,2), (0,3), (1,0), (1,2), (1,3), (2,0), (2,1), (2,3), (3,0), (3,1), (3,2)$. Είναι εμφανές, όμως ότι οι όροι που αντιστοιχούν στους κόμβους (i, j) και (j, i) είναι ίσοι άρα:

$$C = 2 \left(\underbrace{\left(1 - \frac{1}{4}\right)}_{(0,1)} + \underbrace{\left(0 - \frac{1}{4}\right)}_{(0,2)} + \underbrace{\left(1 - \frac{1}{4}\right)}_{(0,3)} + \underbrace{\left(1 - \frac{1}{4}\right)}_{(1,2)} + \underbrace{\left(0 - \frac{1}{4}\right)}_{(1,3)} + \underbrace{\left(1 - \frac{1}{4}\right)}_{(2,3)} \right)$$
$$= 2 \left(4 - 6 \frac{1}{4} \right)$$
$$= 5$$

Τελικά:

$$Q = \frac{1}{16}(5 + 5) = \frac{5}{8}$$

Αυτή είναι και η διαμέριση με το μέγιστο modularity score.

Θέμα 5

A.

(α) MSEIR (β) SIRD (γ) SEIS (δ) SIR

B.

$$
\begin{aligned}
\frac{dS(t)}{dt} &= aM(t) - \lambda S(t) + \gamma \frac{S(t)}{N} I(t) \\
\frac{dE(t)}{dt} &= \lambda S(t) - \beta E(t) \\
\frac{dI(t)}{dt} &= \beta E(t) - \kappa I(t) - \gamma \frac{S(t)}{N} I(t) \\
\frac{dR(t)}{dt} &= \kappa I(t) \\
\frac{dM(t)}{dt} &= -\alpha M(t)
\end{aligned}
$$

Παρατηρούμε ότι:

$$
\frac{d(S+E+I+R+M)}{dt} = 0,
$$

άρα το άθροισμα $S(t) + E(t) + I(t) + R(t) + M(t) = N$ είναι σταθερό ενώ θα ισχύουν και τα παρακάτω:

$$
\begin{aligned}
&S(t), E(t), I(t), R(t), M(t) \geq 0 \\
&S(0) + E(0) + I(0) + R(0) + M(0) = N
\end{aligned}
$$