Large and Social Networks

1ο Σετ Ασκήσεων

Σταυρόπουλος Αλέξανδρος Ανδρέας
2019030109

Διδάσκων:
Θρασύβουλος Σπυρόπουλος

Υπεύθυνος εργαστηρίου:
-

ΗΜΜΥ
Πολυτεχνείο Κρήτης
Χειμερινό εξάμηνο 2022-2023

Πίνακας Περιεχομένων

Άσκηση 1

Ερώτημα 1

Δεδομένου του ότι η Alice έχει 50 φίλους και κάθε μέρα στέλνει ένα email το οποίο είναι μολυσμένο, η πιθανότητα να σταλεί σε ένα συγκεκριμένο φίλο της το mail είναι ίση με $\frac{1}{50}$. Η πιθανότητα αυτή είναι ίση για κάθε έναν από τους φίλους της και δεν επηρεάζεται από το αν στο παρελθόν ο εκάστοτε φίλος έχει λάβει μολυσμένο email.

Η πιθανότητα να σταλεί email σε φίλο ο οποίος προηγουμένως δεν έχει μολυνθεί αλλάζει με την πάροδο του χρόνου και γίνεται όλο και μικρότερη. Πιο συγκεκριμένα, την πρώτη φορά που στέλνεται ένα email, είναι σίγουρο πως θα σταλεί σε φίλο ο οποίος δεν έχει λάβει προηγουμένως, δηλαδή $p_i = 1$. Την δεύτερη φορά, για να μολυνθεί κάποιος νέος φίλος πρέπει να γίνει επιλογή μεταξύ των υπόλοιπων 49, το οποίο σημαίνει πως η πιθανότητα να μολυνθεί δεύτερο άτομο είναι ίση με $p_i = \frac{49}{50}$. Αντίστοιχα, για να μολυνθεί και τρίτο διαφορετικό άτομο η πιθανότητα είναι ίση με $p_i = \frac{48}{50}$ κ.ο.κ.

Αναπτύσσοντας την παραπάνω σκέψη, η πιθανότητα για την μόλυνση του i-οστού νέου ατόμου προκύπτει ως εξής:

$$p_i = \frac{N-i}{N} \tag{1}$$

όπου N ο αριθμός των φίλων της Alice.

Η αναμενόμενη τιμή για την μόλυνση του i-οστού ατόμου προκύπτει ως εξής:

$$E\left[i\right] = \frac{1}{p_i} = \frac{N}{N-i} \tag{2}$$

Ελέγχοντας τις τιμές της αναμενόμενης τιμής παρατηρείται το εξής:

$$\begin{aligned} E\left[n\right] = E\left[n-1\right] + \frac{1}{p_{n-1}} &\Rightarrow E\left[n\right] = E\left[n-2\right] + \frac{1}{p_{n-2}} + \frac{1}{p_{n-1}} \\ &\vdots \\ &\Rightarrow E\left[n\right] = \frac{1}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_{n-2}} + \frac{1}{p_{n-1}} \end{aligned}$$

και αντικαθιστώντας την σχέση (2) η παραπάνω σχέση μπορεί να μετασχηματιστεί ως εξής:

$$\begin{aligned} E\left[n\right] = \frac{1}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_{n-2}} + \frac{1}{p_{n-1}} &\Rightarrow E\left[n\right] = \frac{N}{N-0} + \frac{N}{N-1} + \cdots + \frac{N}{N-(N-2)} + \frac{N}{N-(N-1)} \Rightarrow \\ &\Rightarrow E\left[n\right] = N \cdot \left(\frac{1}{N-0} + \frac{1}{N-1} + \cdots + \frac{1}{N-(N-2)} + \frac{1}{N-(N-1)}\right) \Rightarrow \\ &\Rightarrow E\left[n\right] = N \cdot \left(\frac{1}{N} + \frac{1}{N-1} + \cdots + \frac{1}{2} + 1\right) \Rightarrow \\ &\Rightarrow E\left[n\right] = N \cdot \sum_{n=0}^{N-1} \frac{1}{n-1} \end{aligned} \tag{3}$$

Αντικαθιστώντας τον αριθμό των φίλων της Alice, δηλαδή N = 50 προκύπτει η εξής αναμενόμενη τιμή:

$$E\left[50\right] = 50 \cdot \sum_{n=0}^{49} \frac{1}{n-1} = 224.96 \approx 225 \text{ μέρες}$$

δηλαδή απαιτούνται περίπου 225 ημέρες ώστε να μολυνθεί κάθε ένας από τους φίλους της Alice.

Ερώτημα 2

Όπως προαναφέρθηκε στο προηγούμενο ερώτημα, η πιθανότητα να μολυνθεί ένας φίλος της Alice κάθε μέρα, είναι ίση με $\frac{1}{50}$. Γνωρίζοντας αυτό, η πιθανότητα να μην μολυνθεί ένας φίλος της Alice μία μέρα είναι ίση με $1 - \frac{1}{50} = \frac{49}{50}$ οπότε σε διάστημα 3 μηνών, δηλαδή 90 ημερών, η πιθανότητα να μην μολυνθεί ένας φίλος της Alice είναι ίση με:

$$P_{90} = \sum_{n=0}^{90} \left(1 - \frac{1}{50}\right) = \left(\frac{49}{50}\right)^{90} \tag{4}$$

οπότε η πιθανότητα να μολυνθεί ένας φίλος της Alice στις επόμενες 90 μέρες είναι ίση με:

$$1 - P_{90} = 1 - \left(\frac{49}{50}\right)^{90} \tag{5}$$

Η πιθανότητα $\bar{P_{90}}$ αντιστοιχεί στην πιθανότητα να μολυνθεί ένας φίλος της Alice στις πρώτες 90 μέρες, και έτσι εφόσον η Alice έχει 50 φίλους, για την εύρεση του αριθμού των φίλων της που θα έχουν μολυνθεί στις πρώτες 90 μέρες αρκεί απλά να πολλαπλασιαστεί η πιθανότητα της σχέσης (5 με τον αριθμό των φίλων της:

$$\bar{P_{90}} = 50 \cdot \left(1 - \left(\frac{49}{50}\right)^{90}\right) = 41.884 \approx 42 \text{ φίλοι} \tag{6}$$

Άσκηση 2

Ερώτημα 1

Σύμφωνα με τα δεδομένα της εκφώνησης η Discrete Time Markov Chain σχεδιάζεται ως εξής:

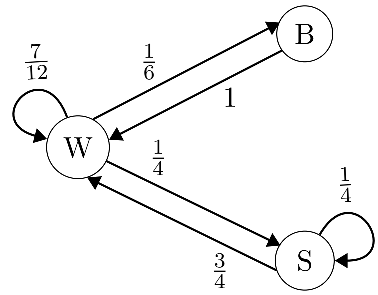

Figure 1: Discrete Markov Chain

Ερώτημα 2

Για να είναι εργοτική η DTMC του figure 3 είναι απαραίτητο να ισχύουν οι εξής προϋποθέσεις:

1. Recurrent (Ελληνικά)
2. Απεριοδική
3. All nodes communicating to each other

Recurrent

Η DTMC είναι Recurrent εφόσον δεν υπάρχει κάποιο state το οποίο είναι transient, δηλαδή δεν υπάρχει κάποιο state από το οποίο αν γίνει μετάβαση από αυτό, δεν είναι εφικτή η επιστροφή σε αυτό.

Απεριοδικότητα

Η DTMC είναι Απεριοδική εφόσον δεν υπάρχει κάποιο state το οποίο είναι περιοδικό, δηλαδή δεν υπάρχει κάποιο state από το οποίο όλα τα πιθανά μονοπάτια που ξεκινούν και καταλήγουν σε αυτό να είναι πολλαπλάσια του ίδιου αριθμού Κ (Κ>1). Αυτό συμβαίνει καθώς υπάρχουν self loops σε 2 από τα 3 states με αποτέλεσμα η μετάβαση να μπορεί να μεγαλώσει αρκετά αξιοποιώντας τα.

π.χ. για το state B, η μετάβαση $B \to W \to B$ έχει ελάχιστο μήκος διαδρομής, ίσο με 2, όμως αντίστοιχα μπορεί να εκτελεστεί και η διαδρομή $B \to W \to W \to B$ η οποία έχει μήκος ίσο με 3.

All nodes connect with each other

Από το figure (1) είναι εμφανές πως όλοι οι κόμβοι επικοινωνούν μεταξύ τους, δηλαδή από κάθε ένα κόμβο, μπορεί να γίνει μετάβαση σε έναν άλλο.

Άρα, πληρούνται όλες οι προϋποθέσεις και ως αποτέλεσμα η DTMC είναι εργοτική

Ερώτημα 3

Για να είναι χρονικά-αναστρέψιμη η DTMC πρέπει να ισχύει το local balance. Εφαρμόζοντας local balance σε κάθε ζευγάρι κόμβων προκύπτουν οι εξής σχέσεις:

$$P_W \cdot \frac{1}{6} = P_B \cdot 1 \Rightarrow P_B = P_W \cdot \frac{1}{6} \tag{7}$$
$$P_W \cdot \frac{1}{4} = P_S \cdot \frac{3}{4} \Rightarrow P_S = P_W \cdot \frac{1}{3} \tag{8}$$

Ακόμα, πρέπει να ισχύει η εξής σχέση:

$$P_W + P_B + P_S = 1$$

οπότε αντικαθιστώντας τις σχέσεις (7), (8) προκύπτουν τα εξής αποτελέσματα:

$$\begin{aligned} P_W + P_B + P_S = 1 \Rightarrow P_W + P_W \cdot \frac{1}{6} + P_W \cdot \frac{1}{3} = 1 \Rightarrow \\ \Rightarrow P_W \left(\frac{6}{6} + \frac{1}{6} + \frac{2}{6}\right) = 1 \Rightarrow \\ \Rightarrow P_W \left(\frac{3}{2}\right) = 1 \Rightarrow \\ \Rightarrow P_W = \frac{2}{3} \end{aligned} \tag{9}$$

και από τις σχέσεις (7), (8) οι τιμές των P_S και P_B προκύπτουν ως εξής:

$$\xRightarrow[(7)]{(9)} P_B = P_W \cdot \frac{1}{6} = \frac{2}{3} \cdot \frac{1}{6} \Rightarrow P_B = \frac{1}{9} \tag{10}$$
$$\xRightarrow[(8)]{(9)} P_S = P_W \cdot \frac{1}{3} = \frac{2}{3} \cdot \frac{1}{3} \Rightarrow P_S = \frac{2}{9} \tag{11}$$

Εφόσον οι τιμές των P_W, P_B, P_S είναι μικρότερες του 1 και το άθροισμά τους είναι ίσο με 1, τότε ισχύει το local balance και κατ' εξοχήν η DTMC είναι χρονικά-αναστρέψιμη.

Ερώτημα 4

Ο χρόνος για τον οποίο θα λειτουργεί το data center είναι ο χρόνος που θα βρίσκεται στο state W, το οποίο σύμφωνα με το προηγούμενο ερώτημα, η πιθανότητα να βρίσκεται στο P_W είναι ίσο με $\frac{2}{3}$, δηλαδή το data center θα είναι ενεργό το 66,6% του χρόνου.

Ερώτημα 5

Η πιθανότητα βρεθούμε σε backhoe είναι ίση με την πιθανότητα μετάβασης από το working state στο backhoe state. Η πιθανότητα αυτή είναι ίση με $\frac{1}{9}$ οπότε ο αναμενόμενος χρόνος μεταξύ σφάλματος λόγω backhoe είναι ίσος $T = \frac{1}{\frac{1}{9}} = 9$ ημέρες.

Άσκηση 3

Ερώτημα 1

Αναπαριστούμε τα states ως 3 αριθμοί, ο πρώτος είναι η τωρινή σελίδα και οι άλλοι 2 αριθμοί είναι οι σελίδες που είναι cached οπότε προκύπτουν 5 πιθανά states: (1,1,2) (2,1,2) (1,1,3) (3,1,3) (3,2,3) (2,2,3) και μέσω αυτών κατασκευάζεται ο πίνακας μετάβασης:

Η λύση του προβλήματος προκύπτει λύνοντας τις παρακάτω σχέσεις:

$$\begin{aligned}
&\pi_{(3,2,3)} = \pi_{(2,2,3)}(1-y) + \pi_{(1,1,2)}(1-x) + \pi_{(2,1,2)}(1-y)\\
&\pi_{(2,2,3)} = \pi_{(3,2,3)} + \pi_{(3,1,3)} + \pi_{(1,1,3)}x\\
&\pi_{(1,1,3)} = \pi_{(2,2,3)}y\\
&\pi_{(3,1,3)} = \pi_{(1,1,3)}(1-x)\\
&\pi_{(2,1,2)} = \pi_{(1,1,2)}x\\
&1 = \pi_{(3,2,3)} + \pi_{(2,2,3)} + \pi_{(1,1,3)} + \pi_{(3,1,3)} + \pi_{(2,1,2)} + \pi_{(1,1,2)}
\end{aligned}$$

οπότε λύνοντας τις παραπάνω σχέσει προκύπτουν οι παρακάτω τιμές:

$$\begin{aligned}
\pi_{(2,1,2)} &= 0\\
\pi_{(1,1,2)} &= 0\\
\pi_{(2,2,3)} &= \frac{1}{2+y-xy}\\
\pi_{(3,1,3)} &= \frac{(1-x)y}{2+y-xy}\\
\pi_{(3,2,3)} &= \frac{1-y}{2+y-xy}\\
\pi_{(1,1,3)} &= \frac{y}{2+y-xy}
\end{aligned}$$

Ερώτημα 2

Για να βρεθεί ο αριθμός των request προς σελίδες που είναι ήδη cashed, έτσι η ζητούμενη τιμή είναι ίση με το άθροισμα όλων των request προς cashed σελίδες επί την πιθανότητα μετάβασης σε αυτή.

Άσκηση 4

Ερώτημα 1

Σύμφωνα με τα δεδομένα κάθε πίνακα, είναι εμφανές πως και για το A και για το B η πιθανότητα να βρεθούν σε ένα οποιοδήποτε κανάλι ξεκινώντας από οποιοδήποτε κανάλι, είναι ίση $\frac{1}{3}$. Έτσι, είναι προφανές πως για να βρεθούν στο ίδιο κανάλι απαιτούνται $\frac{1}{\frac{1}{3}}$ βήματα (δηλαδή 3 βήματα) κατά μέσο όρο δηλαδή αναμένεται να υπάρχει "σύγκρουση" αν 3 βήματα κατά μέσο όρο.

Ερώτημα 2

Στο ερώτημα αυτό οι πίνακες δεν έχουν ίση πιθανότητα ($\frac{1}{3}$) και έτσι για την επίλυση το προβλήματος χρησιμοποιείται και πάλι Markov chain. Πιο συγκεκριμένα, η πιθανότητα να μεταδώσει σε νέο κανάλι κάθε μεταβλητή βασίζεται στο κανάλι που βρίσκεται τώρα και την πιθανότητα μετάβασης προς το νέο κανάλι.

Είναι προφανές πως τα states (1,1), (2,2) (3,3) είναι absorbing καθώς όταν βρεθούν σε αυτά, τότε εκπέμπουν στο ίδιο κανάλι. Έτσι, ξεκινώντας από την κατάσταση ίδια αρχική κατάσταση (1,3) κατασκευάζεται ο παρακάτω πίνακας:

	(1,1)	(2,2)	(3,3)	(1,2)	(1,3)	(2,1)	(2,3)	(3,1)	(3,2)
(1,1)	1	0	0	0	0	0	0	0	0
(2,2)	0	1	0	0	0	0	0	0	0
(3,3)	0	0	1	0	0	0	0	0	0
(1,2)	0.2 * 0.2	0.6 * 0.5	0.2* 0.3	0.2 * 0.5	0.2* 0.3	0.6 * 0.2	0.6 * 0.3	0.2 * 0.2	0.2 * 0.5
(1,3)	0.2 * 0.3	0.6 *0.3	0.2 * 0.4	0.2 *0.3	0.2 * 0.4	0.6 * 0.3	0.6 * 0.4	0.2 * 0.3	0.2 * 0.3
(2,1)	0.4 *0.4	0.3 * 0.3	0.3 * 0.3	0.4 * 0.3	0.4 * 0.3	0.3 * 0.4	0.3 * 0.3	0.3 *0.4	0.3 * 0.3
(2,3)	0.4 * 0.3	0.3 * 0.3	0.3 * 0.4	0.4 * 0.3	0.4 * 0.4	0.3 * 0.3	0.3 * 0.4	0.3 * 0.3	0.3 * 0.3
(3,1)	0.5 * 0.4	0.1 * 0.3	0.4 * 0.3	0.5 * 0.3	0.5 * 0.3	0.1 * 0.4	0.1 * 0.3	0.4 * 0.4	0.4 * 0.3
(3,2)	0.5 * 0.2	0.1 * 0.5	0.4 * 0.3	0.5 * 0.5	0.5 * 0.3	0.1 * 0.2	0.1 * 0.3	0.4 * 0.2	0.4 * 0.5

Οπότε έχουμε:

1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0.04	0.3	0.06	0.1	0.06	0.12	0.18	0.04	0.1
0.06	0.18	0.08	0.06	0.08	0.18	0.24	0.06	0.06
0.16	0.09	0.09	0.12	0.12	0.12	0.09	0.12	0.09
0.12	0.09	0.12	0.12	0.16	0.09	0.12	0.09	0.09
0.2	0.03	0.12	0.15	0.15	0.04	0.03	0.15	0.12
0.1	0.05	0.12	0.25	0.15	0.02	0.03	0.08	0.2

όπου ο παραπάνω πίνακας ακολουθεί την μορφή $P = \begin{bmatrix} Q & R \\ 0 & I_r \end{bmatrix}$ με Q να είναι ο εξής πίνακας:

0.1	0.06	0.12	0.18	0.04	0.1
0.06	0.08	0.18	0.24	0.06	0.06
0.12	0.12	0.12	0.09	0.12	0.09
0.12	0.16	0.09	0.12	0.09	0.09
0.15	0.15	0.04	0.03	0.15	0.12
0.25	0.15	0.02	0.03	0.08	0.2

Τελίκα, η απάντηση του προβλήματος δίνεται από τον εξής πίνακα:

$$N = (1-Q)^{-1} = \begin{bmatrix} 1.335 & 0.273 & 0.294 & 0.394 & 0.195 & 0.294 \\ 0.319 & 1.325 & 0.381 & 0.483 & 0.232 & 0.273 \\ 0.374 & 0.349 & 1.312 & 0.326 & 0.293 & 0.301 \\ 0.375 & 0.392 & 0.290 & 1.369 & 0.262 & 0.302 \\ 0.404 & 0.372 & 0.226 & 0.264 & 1.328 & 0.332 \\ 0.541 & 0.394 & 0.229 & 0.299 & 0.256 & 1.445 \end{bmatrix} \tag{12}$$

και προσθέτωντας τις τιμές της δεύτερης στήλης που αντιστοιχεί στο (1,3) προκύπτει η ζητούμενη απάντηση = 3,105