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

Εισαγωγή

Στην πρώτη άσκηση ζητείται η υλοποίηση των αλγορίθμων ε-Greedy και Upper Confidence Bound, οι οποίοι επιτυγχάνουν ισορροπία μεταξύ exploration και exploitation στο γνωστό πρόβλημα κουλοχέρηδων (Bandits Problem). Εφόσον υλοποιήθηκε κάθε αλγόριθμος, δόθηκε ως όρισμα το ίδιο σύνολο από bandits ώστε να εκτελεστούν για συγκεκριμένο αριθμό γύρων (Ορίζοντας T) και επίσης επιλέχθηκε ένας από τους bandits με βάση τον γινόμενο μεταξύ reward και πιθανότητας επιτυχίας. Η επίδοση κάθε αλγορίθμου βασίζεται στο regret το οποίο είναι η διαφορά μεταξύ του σκορ του αρχικά επιλεγμένου bandit και του σκορ που μάζεψε ο κάθε αλγόριθμος.

ε-Greedy

Στον αλγόριθμο ε-Greedy στην περίπτωση exploration επιλέγεται τυχαίο χέρι ανεξάρτητα των επιδόσεών του έως τώρα ενώ στην περίπτωση exploitation επιλέγεται το χέρι με την καλύτερη επίδοση. Η επίδοση κάθε χεριού ορίζεται μέσω του συντελεστή μ ο οποίος είναι ίσος με τον πηλίκο μεταξύ του σκορ που έχει μαζέψει το εκάστοτε χέρι προς τον αριθμό των φορών που έχει επιλεχθεί. Το σκορ προκύπτει ως το γινόμενο μεταξύ reward και διωνυμικής πιθανότητας κάθε χεριού και υπολογίζεται σε κάθε γύρο ξεχωριστά.

Η επιλογή μεταξύ exploitation και exploitation γίνεται με χρήση της μεταβλητής epsilon η οποία ορίζεται ως εξής:

$$epsilon = (t)^{-\frac{1}{3}} \cdot (k \cdot \log(t))^{\frac{1}{3}}$$

Σε κάθε γύρο, με πιθανότητα epsilon γίνεται explore ενώ με πιθανότητα 1 - epsilon γίνεται exploit και λαμβάνοντας υπόψιν τον φθίνον ρυθμό της μεταβλητής epsilon, μπορεί εύκολα να αποδειχθεί πως όσο αυξάνονται οι γύροι τόσο πιο πιθανό είναι να γίνει exploitation χρησιμοποιώντας το καλύτερο χέρι ενώ ταυτόχρονα τόσο πιο απίθανο να επιλεγεί κάποιο τυχαίο (exploration). Η υλοποίηση παρουσιάζει convergence rate ίσο με:

$$O\left((t)^{\frac{2}{3}} \cdot (K \cdot \log(t))^{\frac{1}{3}}\right) \tag{1}$$

Upper Confidence Bound

Στον αλγόριθμο Upper Confidence Bound (UCB) η επιλογή χεριού γίνεται με βάση τον συντελεστή ubc:

$$ucb = \mu + \sqrt{\frac{\log T}{Q}}$$

όπου μ ο συντελεστής επίδοσης κάθε χεριού, T ο ορίζοντας και Q οι φορές που έχει επιλεχθεί το αντίστοιχο χέρι.

Σε κάθε γύρο επιλέγεται το χέρι το οποίο παρουσιάζει μεγαλύτερο συντελεστή ucb κάτι το οποίο προκύπτει είτε λόγο καλής επίδοσης (συντελεστής μ) είτε επειδή έχουν περάσει πολλοί γύροι που δεν έχει επιλεχθεί το εκάστοτε χέρι. Αυτό έχει ως αποτέλεσμα, το exploration πρακτικά να μην σταματάει ποτέ καθώς ακόμα και για μεγάλο αριθμό γύρου, ανά διαστήματα επιλέγεται διαφορετικό χέρι και αξιολογείται εκ νέου το performance του κάτι το οποίο στον ε-Greedy είναι σχεδόν απίθανο να συμβεί. Η υλοποίηση παρουσιάζει convergence rate ίσο με:

$$O\left(\sqrt{K \cdot T \cdot \log T}\right) \tag{2}$$

Σύγκριση επιδόσεων

Εξετάζοντας τα convergence rate κάθε αλγορίθμου αναμένεται η επίδοση του UCB να είναι καλύτερη σε σχέση με του ε-Greedy. Για την επαλήθευση αυτής της εκτίμησης, έγινε σύγκριση των επιδόσεων μεταξύ των δύο αλγορίθμων για το ίδιο σετ bandits σε κάθε τεστ. Αρχικά, για 10 bandits και ορίζοντα μεγέθους 1000, τα αποτελέσματα που εξάχθηκαν δεν συναδουν πάντα με την εκτίμηση. Συχνότερη περίπτωση αποτελεί, όπως ήταν αναμενόμενο, ο UCB να παρουσιαζει καλύτερο performance δηλαδή μικρότερο regret (fig:1), ωστόσο, υπήρξαν σπάνιες περιπτώσεις όπου ο ε-Greedy παρουσίασε καλύτερη επίδοση (fig:2) κάτι το οποίο είναι πιθανό να συμβεί καθώς το σκορ κάθε αλγορίθμου υπολογίζεται ξεχωριστά οπότε για τυχαίους λόγους ο ένας αλγόριθμος να είναι πολύ πιο "τυχερός" από τον άλλο.

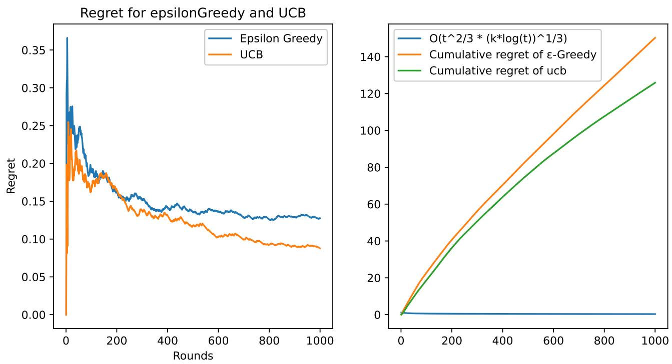

Figure 1: Lower regret for UCB

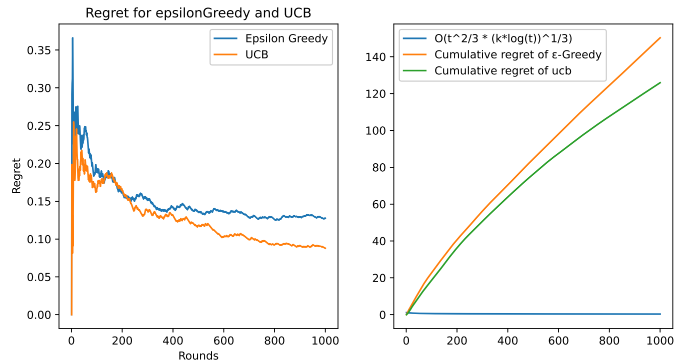

Figure 2: Lower regret for ε-Greedy

Για την καλύτερη επιβεβαίωση των αποτελεσμάτων, αρχικά αυξήθηκε η τιμή ορίζοντα ($T = 10000$) ενώ ο αριθμός των bandit παρέμεινε σταθερός. Στην περίπτωση αυτή ο UCB πάντα παρουσιάζει καλύτερη επίδοση, δηλαδή μικρότερο regret, σε σχέση με τον ε-Greedy κάτι το οποίο επιβεβαιώνει την αρχική εκτίμηση

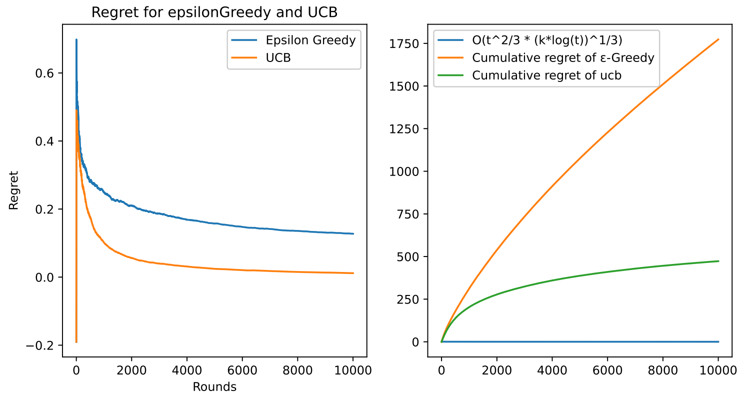

Figure 3: Lower regret for ε-Greedy

Παρατηρώντας τις γραφικές κάθε περίπτωσης πέραν την καλύτερης επίδοσης του UCB παρατηρείται πως και οι δύο αλγόριθμοι επιτυγχάνουν sublinear πολυπλοκότητα δεδομένου του οτι ο ρυθμός αύξησης του regret μειώνεται όσο αυξάνονται οι γύροι.

- πες οτι ο άλλος είναι πιο απλός και δεν βγάζει τόσο κακά αποτελέσματα

Άσκηση 2

Ερώτημα 1

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

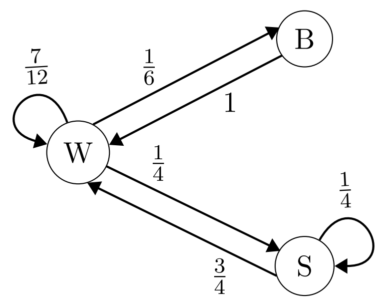

Figure 4: Discrete Markov Chain

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 (4) 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:

$$P_W \cdot \frac{1}{6} = P_B \cdot 1 \Rightarrow P_B = P_W \cdot \frac{1}{6} \tag{3}$$

$$P_W \cdot \frac{1}{4} = P_S \cdot \frac{3}{4} \Rightarrow P_S = P_W \cdot \frac{1}{3} \tag{4}$$

Ακόμα, πρέπει να ισχύει η εξής σχέση:

$$P_W + P_B + P_S = 1$$

οπότε αντικαθιστώντας τις σχέσεις (3), (4) προκύπτουν τα εξής αποτελέσματα:

$$\begin{aligned} P_W + P_B + P_S = 1 \Rightarrow P_W + P_W \cdot \frac{1}{6} + P_W \cdot \frac{1}{3} = 1 \Rightarrow \\ \Rightarrow P_W \left(\frac{6}{6} + \frac{1}{6} + \frac{2}{6} \right) = 1 \Rightarrow \\ \Rightarrow P_W \left(\frac{3}{2} \right) = 1 \Rightarrow \\ \Rightarrow P_W = \frac{2}{3} \end{aligned} \tag{5}$$

και από τις σχέσεις (3), (4) οι τιμές των P_S και P_B προκύπτουν ως εξής:

$$\xRightarrow[(3)]{(5)} P_B = P_W \cdot \frac{1}{6} = \frac{2}{3} \cdot \frac{1}{6} \Rightarrow P_B = \frac{1}{9} \tag{6}$$

$$\xRightarrow[(4)]{(5)} P_S = P_W \cdot \frac{1}{3} = \frac{2}{3} \cdot \frac{1}{3} \Rightarrow P_S = \frac{2}{9} \tag{7}$$

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

Άσκηση 3

Ερώτημα 1

Ερώτημα 2

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