

Θεωρία γραφημάτων και εφαρμογές (ICE 8108)

Εξαμηνιαία εργασία

Διδάσκων: Γεώργιος Δρακόπουλος

28 Μαΐου 2025

1 Θεωρητικό μέρος (4 μονάδες)

Στο τμήμα αυτό θα δώσετε απαντήσεις χωρίς να προγραμματίσετε. Εκτός από τυχόν εξισώσεις θα πρέπει να συμπεριλάβετε και βασικά επιχειρήματα οι οποίες είτε δικαιολογούν είτε πλαισιώνουν τις απαντήσεις που θα δώσετε. Απαντήστε σε όλα τα ερωτήματα. Θα χρειαστεί να πάρετε τρεις τουλάχιστον μονάδες από το σκέλος αυτό για να λάβετε προβιβάσιμο βαθμό. Κάθε ερώτημα μετράει για μια μονάδα.

1.1 Κατανομή βαθμών

- Ποιά είναι η πιθανότητα μια τυχαία μεταβλητή X η οποία ακολουθεί την κανονική κατανομή της εξίσωσης (1) με μέση τιμή μ_0 μηδέν και μοναδιαία διασπορά σ_0^2 να πάρει την τιμή μηδέν;

$$f_X(x; \mu_0, \sigma_0^2) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_0)^2}{2\sigma_0^2}\right) \quad (1)$$

- Για ποιούς λόγους η κατανομή Gauss δεν αποτελεί καλή προσέγγιση της κατανομής των βαθμών των κορυφών σε γράφους ανεξάρτητους κλιμάκωσης (scale free graphs);

1.2 Εκθετικό μητρείο γειτνίασης

Αποδείξτε πως στην περίπτωση όπου το μητρείο γειτνίασης \mathbf{A} έχει φάσμα τέτοιο ώστε η αριθμητική και γεωμετρική πολλαπλότητα κάθε ιδιοτιμής να ισούται με την μονάδα το αντίστοιχο εκθετικό μητρείο $e^{\mathbf{A}}$ έχει θετικές ιδιοτιμές.

1.3 Συμπίεση μητρείου γειτνίασης

Ποιές είναι οι διαφορές όταν το μητρείο γειτνίασης ενός γράφου παραγοντοποιείται μέσω της παραγοντοποίησης ιδιζουσών τιμών (SVD) και μέσω του διδιάστατου διακριτού μετασχηματισμού συνημιτόνου (DCT2); Για να είναι ολοκληρωμένη η απάντησή σας, σκεφτείτε κατ' ελάχιστον την ερμηνεία των δύο μεθοδολογιών, τον χώρο τον οποίο καταλαμβάνουν οι αντίστοιχες συμπίεσμένες μορφές, και την πολυπλοκότητα υπολογισμού κάθε μετασχηματισμού. Επιπλέον διαφορές θα προσμετρηθούν θετικά.

1.4 Ταξινόμηση ακεραίων

Έστω πως έχουμε στην διάθεσή μας n διακριτούς αλλά όχι απαραίτητα συνεχόμενους ακέραιους.

- Αν κάθε τέτοιος ακέραιος αποτελεί την κορυφή ενός γράφου, τότε πώς μπορεί να αναχθεί η ταξινόμησή τους σε αναζήτηση συντομότερων διαδρομών; Τι δομή θα πρέπει να έχει αυτός ο γράφος; Στην περίπτωση αυτή, πόση θα είναι η πολυπλοκότητα του αλγορίθμου του Dijkstra και ποιός όρος κυριαρχεί στην πολυπλοκότητα; Πόση είναι η χωρική πολυπλοκότητα του εν λόγω γράφου; Τι παρατηρείτε;
- Αν οι $n!$ δυνατές διατάξεις των ανωτέρω ακεραίων τοποθετηθούν στα φύλλα ενός δυαδικού δέντρου και κάθε ενδιάμεση κορυφή αποτελεί μια σύγκριση μεταξύ δύο ακεραίων, τότε ποιά είναι το ύψος αυτού του δέντρου; Πόσο είναι το μήκος ενός μονοπατιού από την ρίζα προς κάθε φύλλο; Τι παρατηρείτε ως προς την πολυπλοκότητα και ποιά η διαφορά με εκείνη του προηγούμενου υποερωτήματος; Ποιά είναι τώρα η χωρική πολυπλοκότητα; Τι παρατηρείτε;

2 Προγραμματιστικό μέρος (6 μονάδες)

Από το παρόν τμήμα θα πρέπει να επιλέξετε τρία ερωτήματα και να τα υλοποιήσετε σε Python. Σε τυχόν περίπτωση όπου επιλέξετε περισσότερα ερωτήματα, τότε θα προσμετρηθούν εκείνα τα ερωτήματα τα οποία δίνουν την μεγαλύτερη βαθμολογία. Κάθε ερώτημα μετράει για δύο μονάδες.

2.1 Αναζήτηση κορυφών

Υλοποιήστε την αναζήτηση πρώτα κατά βάθος (DFS) και πρώτα κατά πλάτος (BFS) για τον γράφο graph1.csv. Η κορυφή εκκίνησης είναι η πρώτη αριθμητικά. Τι παρατηρείτε; Μπορείτε να φτάσετε σε όλες τις κορυφές από την κορυφή εκκίνησης;

2.2 Συντελεστής συσταδοποίησης

Για τον γράφο graph1.csv υπολογίστε και τυπώστε την κατανομή των τριγώνων και την τιμή του συντελεστή συσταδοποίησης c_i για κάθε μια από τις n κορυφές v_i . Με βάση αυτούς τους συντελεστές, υπολογίστε τον ολικό συντελεστή συσταδοποίησης C του γράφου σύμφωνα με την εξίσωση (2). Σε αυτή t_i και o_i είναι αντίστοιχα ο αριθμός των τριγώνων και ο αριθμός της δεδομένης κορυφής. Για να υπολογίσετε τον ολικό αριθμό ανοιχτών τριγώνων, χρησιμοποιείστε συνδυαστικά επιχειρήματα.

$$C = \frac{1}{n} \sum_{i=1}^n c_i = \frac{1}{n} \sum_{i=1}^n \frac{t_i}{o_i} \quad (2)$$

Επίσης υπολογίστε τον εναλλακτικό συντελεστή συσταδοποίησης C' σύμφωνα με την εξίσωση (3).

$$C' = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n o_i} \quad (3)$$

Τι παρατηρείτε; Ισχύει γενικά η παρατήρησή σας; Μετατρέψτε πρώτα τον γράφο σε μη κατευθυνόμενο για τους υπολογισμούς σας.

2.3 Κεντρικότητα ιδιοδιανύσματος

Για τον γράφο `graph1.csv` υπολογίστε το πρωτεύον ιδιοδιάνυσμα και τυπώστε την κατανομή των στοιχείων του. Ποιά είναι η συσχέτιση με την κατανομή των βαθμών των κορυφών; Επίσης προσαρμόστε με την μέθοδο των ελάχιστων τετραγώνων έναν νόμο εκθέτη της μορφής της εξίσωσης (4) στις δύο αυτές κατανομές.

$$f[k] = \alpha_0 k^{-\gamma_0} \quad (4)$$

Τι εκθέτη γ_0 πήρατε στις δύο περιπτώσεις; Τι παρατηρείτε; Μετατρέψτε πρώτα τον γράφο σε μη κατευθυνόμενο για να κάνετε τους υπολογισμούς σας.

2.4 Συντομότερες διαδρομές

Υλοποιήστε τον αλγόριθμο Dijkstra για τους γράφους `dijkstra1.csv` και `dijkstra2.csv`. Το κόστος ισούται με τον αριθμό των ακμών, επομένως ουσιαστικά κάθε ακμή έχει μοναδιαίο κόστος. Και στις δύο περιπτώσεις η κορυφή εκκίνησης είναι η πρώτη σε σειρά αρίθμησης, και η κορυφή προορισμού είναι αντίστοιχα η τελευταία. Οι κορυφές εκκίνησης και προορισμού υπάρχουν σε σχόλια στην αρχή των σχετικών αρχείων. Τι παρατηρείτε; Πόσο είναι το μήκος της ελάχιστης διαδρομής σε κάθε περίπτωση; Δεδομένου πως ο πρώτος γράφος είναι δέντρο, τι συμπεραίνετε;

2.5 Συνθετικοί γράφοι

Υλοποιείτε τα ακόλουθα μοντέλα συνθετικών γράφων.

- Το μοντέλο $\mathcal{G}_{n,p}$ για n ίσο με 1000 και p ίση με 0.25 για εκατό στιγμιότυπα. Ποιός είναι ο μέσος αριθμός κορυφών και ακμών; Ποιά είναι η διασπορά του;
- Το μοντέλο Barabási-Albert με αρχικό γράφο το `ba.csv` για εκατό φορές μέχρι να διπλασιαστεί ο αριθμός των κορυφών του γράφου. Υπολογίστε και τυπώστε σε γράφημα την μέση πυκνότητα του γράφου σε συνάρτηση με το κάθε βήμα της εκτέλεσης.

Καλή επιτυχία!

Α' Οδηγίες

Πριν ξεκινήσετε την εργασία, διαβάστε προσεκτικά τα ακόλουθα.

- Η εργασία θα εξεταστεί προφορικά σε ημερομηνία η οποία θα ανακοινωθεί. Η εργασία παραμένει η ίδια για την εξεταστική του Σεπτεμβρίου χωρίς διαφορά ή οποιαδήποτε ποινή στην βαθμολογία. Δηλαδή ο τρόπος βαθμολόγησης θα είναι ο ίδιος τον Ιούλιο και τον Σεπτέμβριο.
- Η εργασία γίνεται σε ομάδες μέχρι και τριών ατόμων. Η σύνθεση των ομάδων επαφίεται αποκλειστικά σε εσάς. Οι ομάδες δεν χρειάζεται να δηλωθούν παρά μόνον όταν βγει θα ανακοίνωση για τον προγραμματισμό της προφορικής εξέτασης.

- Όπως προαναφέρθηκε, στο θεωρητικό τμήμα δεν χρειάζεται (και δεν πρέπει) να προγραμματίσετε κάτι. Οι υπολογισμοί θα γίνουν αμιγώς θεωρητικά και θα πρέπει να είναι τεκμηριωμένοι. Μεμονωμένες εξισώσεις χωρίς επιχειρήματα ή προγραμματιστές εντολές ή προσομοιώσεις δεν θεωρούνται αποδεκτές απαντήσεις.
- Στο θεωρητικό μέρος απαντάτε σε όλα τα ερωτήματα, ενώ από το προγραμματιστικό σκέλος επιλέγετε τρία ερωτήματα.
- Για να πάρετε προβιβάσιμο βαθμό, θα πρέπει να πάρετε τουλάχιστον τρεις μονάδες στο θεωρητικό σκέλος.
- Στην περίπτωση όπου επιλέξετε περισσότερα ερωτήματα από όσα σας ζητούνται, τότε προσμετρώνται εκείνα τα οποία δίνουν την μεγαλύτερη βαθμολογία.
- Στο προγραμματιστικό σκέλος οι απαντήσεις θα είναι υλοποιημένες κατά προτίμηση σε Python. Μην παραδώσετε notebook αλλά πλήρη κώδικα.
- Οι γράφοι οι οποίοι αναφέρονται στην εκφώνηση είναι σε αρχεία .csv τα οποία περιέχουν την λίστα γειτνίασης για κάθε κορυφή με μία λίστα σε κάθε γραμμή. Στην πρώτη στήλη κάθε γραμμής υπάρχει η κορυφή την οποία αφορά η συγκεκριμένη λίστα. Ο χαρακτήρας ‘ εισάγει σχόλια.
- Μαζί με την εργασία θα συμπεριλάβετε σε μορφή .csv και τους γράφους οι οποίοι προκύπτουν από την επεξεργασία την οποία θα κάνετε.
- Τα ερωτήματα είναι ισοβαρή. Όταν υπάρχουν υποερωτήματα, τότε αυτά θεωρούνται ισοβαρή μεταξύ τους.
- Η αναφορά της εργασίας θα είναι αποκλειστικά σε μορφή .pdf ή .ps. Αναφορές σε .doc ή .docx δεν θα γίνονται αποδεκτές.
- Η γλώσσα της εργασίας είναι τα ελληνικά.
- Αν υπάρχουν απορίες, τότε αυτές θα πρέπει να αναρτώνται στο e-class του μαθήματος ώστε να διευκολύνονται τυχόν ομάδες οι οποίες με τις ίδιες ή παραπλήσιες απορίες.
- Η εργασία υποβάλλεται αποκλειστικά μέσω του e-class.
- Η χρήση ChatGPT ή άλλων εργαλείων τεχνητής νοημοσύνης απαγορεύεται αυστηρά.
- Σε περίπτωση αντιγραφής θα υπάρξει παραπομπή στα αρμόδια όργανα του τμήματος.
- Ναι, τα ερωτήματα σε κάθε σκέλος είναι ισοβαρή.