



Αλγόριθμοι

Εαρινό Εξάμηνο 2022-2023

Διδάσκων: Δημήτριος Κατσαρός

Βοηθός: Ευαγγελία Φράγκου

Coding project: (σε ομάδες των 2 ατόμων)

Ημέρα ανακοίνωσης: Sunday, March 19, 2023

Προθεσμία παράδοσης: **Τετάρτη, Ιούνιος 28, 2023**



### Περιγραφή προβλήματος

Στην εργασία αυτή θα ασχοληθούμε με συντομότερες διαδρομές σε μη-κατευθυνόμενα (undirected), μη-βεβαρυμένα (unweighted) γραφήματα και την κλιμάκωση του υπολογισμού τους σε γραφήματα καθώς και την δυνατότητά τους να περιγράφουν στοιχεία της τοπολογίας των γραφημάτων.

Το “αφηρημένο” πρόβλημα το οποίο αντιμετωπίζουμε έχει ως εξής: Δίνεται ένα γράφημα (κόμβοι και συνδέσεις) το οποίο περιγράφει την δυνατότητα επικοινωνίας μεταξύ των στοιχείων ενός “οργανισμού”. Οι επικοινωνίες γίνονται με πρωτόκολλα υλοποιημένα με αλγορίθμους συντομοτέρων διαδρομών (shortest paths). Δεν μάς ενδιαφέρουν οι υπόλοιπες τεχνικές λεπτομέρειες των πρωτοκόλλων.

Επιθυμούμε την υλοποίηση δυο ενεργειών:

- [1]. Να μπορούμε να “κρυφακούμε” τις συνομιλίες μεταξύ ζευγών κόμβων εγκαθιστώντας ελάχιστα (ή μόνο ένα, εάν τα οικονομικά μας δεν το επιτρέπουν) κέντρα υποκλοπών.
- [2]. Να εγκαταστήσουμε ελάχιστα (ή μόνο ένα, εάν τα οικονομικά μας δεν το επιτρέπουν) κέντρα διοίκησης που να έχουν την ιδιότητα ότι η επικοινωνία από το κέντρο προς κάθε άλλο κόμβο να είναι όσο το δυνατόν γρηγορότερη.

Αντιμετωπίζουμε το πρακτικό ΕΡΩΤΗΜΑ εάν η λύση του ενός προβλήματος συνεπάγεται και την (σχετικά) καλή λύση και του άλλου προβλήματος. Δηλαδή εάν οι κόμβοι/κόμβος που είναι οι πιο ενδεδειγμένοι για το Πρόβλημα [1] είναι ταυτόχρονα και οι πιο ενδεδειγμένοι για το Πρόβλημα [2].

Για να απαντήσουμε στο ερώτημα το οποίο μάς τέθηκε, ενεργούμε ως εξής:

Σε δεδομένο γράφημα:

- A. Υπολογίζουμε τον συνολικό αριθμό των συντομοτέρων μονοπατιών μεταξύ κάθε ζεύγους κόμβων  $j$  και  $k$  στα οποία μεσολαβεί κάθε κόμβος  $i$  του γραφήματος. Προφανώς αυτός ο αριθμός θα είναι κάποιος ακέραιος αριθμός ή μηδέν.
- B. Διατάσσουμε τους κόμβους του γραφήματος με βάση αυτόν τον αριθμό με διάταξη φθίνουσα ξεκινώντας από εκείνον τον κόμβο με τον μεγαλύτερο τέτοιον αριθμό

- προς εκείνον τον κόμβο με τον μικρότερο τέτοιο αριθμό. Έτσι, θα προκύψει μια διάταξη των IDs των κόμβων του γραφήματος – ας την ονοματίσουμε ως  $R_{\text{ΚΡΥΦ}}$ .
- C. Υπολογίζουμε το μέσο μήκος των συντομοτέρων μονοπατιών μεταξύ ενός συγκεκριμένου κόμβου  $i$  και όλων των υπολοίπων κόμβων του γραφήματος.
- D. Διατάσσουμε τους κόμβους του γραφήματος με βάση αυτό το μέσο μήκος με διάταξη αύξουσα ξεκινώντας από εκείνον τον κόμβο με το μικρότερο μέσο μήκος προς εκείνον τον κόμβο με το μεγαλύτερο μέσο μήκος. Έτσι, θα προκύψει μια διάταξη των IDs των κόμβων του γραφήματος – ας την ονοματίσουμε ως  $R_{\text{ΔΙΟΙΚ}}$ .
- E. Υπολογίζουμε τον Kendall  $\tau$  συντελεστή συσχέτισης μεταξύ των δυο διατάξεων (εάν δεν τον γνωρίζετε ήδη από το μάθημα “Πιθανότητα και Στατιστική” δείτε εδώ [https://en.wikipedia.org/wiki/Kendall\\_rank\\_correlation\\_coefficient](https://en.wikipedia.org/wiki/Kendall_rank_correlation_coefficient)), ο οποίος ορίζεται ως εξής:

$$\tau = 1 - \frac{2 \times \text{disconcordant\_pairs}}{\binom{n}{2}}$$

όπου ως *disconcordant\_pair* σε μια διάταξη ορίζεται ως ένα ζεύγος έστω  $i$  και  $j$  το οποίο ζεύγος στην μια διάταξη το  $i$  εμφανίζεται πριν από το  $j$ , ενώ στην άλλη διάταξη το  $i$  εμφανίζεται μετά το  $j$ . Για παράδειγμα εάν έχουμε δυο διατεταγμένες λίστες στοιχείων (a, b, c, d) και (c, b, a, d) τα *disconcordant pairs* είναι τα (a,b), (a, c) και (b, c), και θα έχουμε  $\tau = 1 - 2 \times 3 / (4 \times 3 / 2) = 1 - 6 / 6 = 0$ .

Ζητείται να υλοποιήσετε τα παραπάνω βήματα A-E πάνω σε τρία γραφήματα, τα οποία διαφέρουν σε μέγεθος και είναι τα εξής:

- i. [Karate club network αποτελούμενο από 34 κόμβους και 78 συνδέσεις] <https://github.com/freditation/karate-club/blob/master/karate.edgelist>
- ii. [Το power grid των δυτικών πολιτειών των Η.Π.Α. αποτελούμενο από 4941 κόμβους και 6594 συνδέσεις – μια σύνδεση αναπαριστά μια γραμμή διανομής ενέργειας, και ένα κόμβος είναι είτε generator, ένας transformator ή ένας substation.] <https://courses.e-ce.uth.gr/ECE216/noexternalweb/powergrid.txt>
- iii. [Paul Erdos συν-συγγραφείς αποτελούμενο από 6927 κόμβους και 11850 συνδέσεις] [https://courses.e-ce.uth.gr/ECE216/noexternalweb/erdos\\_coauthors.txt](https://courses.e-ce.uth.gr/ECE216/noexternalweb/erdos_coauthors.txt)

Για το κάθε γράφημα θα μετρήσετε τον χρόνο εκτέλεσης για την εκτέλεση ενός “αλγορίθμου” και θα φτιάξετε ένα σχεδιάγραμμα επίδοσης, όπου στον άξονα  $x$  θα βρίσκεται το μέγεθος σε κόμβους του γραφήματος και στον άξονα  $y$  ο χρόνος εκτέλεσης. Ανάλογα και όταν ο  $x$  άξονας είναι ο αριθμός των συνδέσεων.

Συνολικά ζητούνται τα εξής:

- ΑΛΓ-[1]. Ψευδοκώδικας και κώδικας για τον αλγόριθμο για τον υπολογισμό των συντομοτέρων μονοπατιών από κάθε κόμβο προς κάθε άλλον κόμβο.
- ΑΛΓ-[2]. Κώδικας για τον υπολογισμό για κάθε κόμβο σε πόσα shortest paths μεσολαβεί.
- ΑΛΓ-[3]. Κώδικας για τον υπολογισμό του μέσου μήκους των συντομοτέρων μονοπατιών από κάθε κόμβο προς όλους τους άλλους.
- ΑΛΓ-[4]. Ψευδοκώδικας και κώδικας για τον υπολογισμό των *disconcordant pairs*

Ως παραδοτέο ζητείται να γράψετε μια αναφορά με τουλάχιστον τα εξής:

- ❑ Ψευδοκώδικες για τους ΑΛΓ[1]-ΑΛΓ-[4], καθώς και την υπολογιστική τους πολυπλοκότητα (ποια είναι και πώς προκύπτει).
- ❑ Παρουσίαση και επεξήγηση των δομών δεδομένων που χρησιμοποιήθηκαν σε κάθε αλγόριθμο.

- ❑ Plot της επίδοσης καθενός εκ των ΑΛΓ[1]-ΑΛΓ-[4] για τα γραφήματα εισόδου με τη μορφή του plot όπως εξηγήθηκε παραπάνω.
- ❑ Πίνακας με τους Kendall τ για κάθε γράφημα εισόδου.
- ❑ Απάντηση στο ΕΡΩΤΗΜΑ που τέθηκε αρχικά και επιχειρηματολογία επί της απάντησης.

---

#### Χρηστικές πληροφορίες:

- Η προθεσμία παράδοσης είναι αυστηρή.
  - Είναι όμως δυνατή η παροχή παράτασης μέχρι 4 ημέρες, αλλά μόνο εφόσον δώσει ο διδάσκων την έγκρισή του.
  - Φροντίστε να ζητήσετε εγκαίρως την παράταση, και όχι την τελευταία ημέρα προ της εκπνοής της προθεσμίας.
  - Αυτή η παράταση στοιχίζει 15% ποινή στον τελικό βαθμό της.
- Η παράδοση του κώδικα της εργασίας και της αναφοράς γίνεται με email με παραλήπτες τους: [dkatsar@uth.gr](mailto:dkatsar@uth.gr) **KAI** [efragkou@uth.gr](mailto:efragkou@uth.gr).
- Το subject του μηνύματος πρέπει να είναι **αυστηρά:** ECE216-Project: AEMx-AEMy (όπου στην θέση των AEMx-AEMy βάλτε τα AEM των μελών της ομάδας χωρισμένα με hyphen).
- Αποτυχία να αποσταλεί και στους δυο παραλήπτες με το καθορισμένο subject πιθανότητα θα σημάνει μη βαθμολόγηση της εργασίας.

#### Ερμηνεία συμβόλου:



Απαιτεί την ανάπτυξη κώδικα σε C/C++, JAVA, Python, κ.τ.λ. **Εάν χρησιμοποιήσετε έτοιμο κώδικα από κάποια πηγή απαιτείται να δηλώσετε την πηγή, καθώς και σε ποια σημεία του project τον χρησιμοποιήσατε.**