

Θέματα Διπλωματικών Εργασιών Χειμερινό Εξάμηνο Ακαδημαϊκού Έτους 2021-2022

- Αν σας ενδιαφέρει κάποιο από τα θέματα στείλτε μου email **μέχρι και τις 24.9.2021** για εκπόνηση διπλωματικής το **χειμερινό εξάμηνο**. Όσοι ενδιαφέρονται για το εαρινό εξάμηνο μπορεί να επικοινωνήσουν μαζί μου και αργότερα.

Διπλωματική 1: Δίκαιος τυχαίος περίπατος (fair random walk)

Ο τυχαίος περίπατος είναι ένας κλασικός τρόπος διάσχισης ενός γράφου. Με απλά λόγια, ξεκινάμε από ένα τυχαίο κόμβο, επισκεπτόμαστε τυχαία κάποιο γείτονα του και μετά τυχαία κάποιο γείτονά του γείτονα κοκ. Οι τυχαίοι περίπατοι αποτελούν βασικό κομμάτι πολλών αλγορίθμων όπως ο PageRank και μιας σημαντικής κατηγορίας graph embedding τεχνικών (τεχνικών μηχανικής μάθησης για αναπαράσταση κόμβων ως διανύσματα).

Σε αυτή τη διπλωματική θα μελετηθούν δίκαιες εκδοχές των τυχαίων περιπάτων. Συγκεκριμένα, σε πολλά δίκτυα (ιδιαίτερα κοινωνικά δίκτυα, όπου οι κόμβοι του δικτύου είναι πρόσωπα και οι συνδέσεις απεικονίζουν σχέσεις, όπως φιλία, συνεργασία, επικοινωνία κλπ), οι κόμβοι χαρακτηρίζονται από ευαίσθητα χαρακτηριστικά, όπως η ηλικία, το γένος, ή το φύλλο τους. Σε τέτοιες περιπτώσεις θα θέλαμε ο τυχαίος περίπατος να είναι δίκαιος, δηλαδή, δισυστημικά, θα θέλαμε να επισκέπτεται κόμβους με όλα τα χαρακτηριστικά (π.χ., και άντρες και γυναίκες).

Σε προηγούμενη δουλειά μας (στην εργασία [1]) έχουμε προτείνει μια δίκαιη εκδοχή του τυχαίου περιπάτου σε συνδυασμό με το PageRank [2].

Ο αλγόριθμος PageRank είναι ίσως ο πιο γνωστός αλγόριθμος υπολογισμού της σημαντικότητας ενός κόμβου στο δίκτυο και ήταν ο βασικός αλγόριθμος πίσω από την επιτυχία της Google. **Καθώς τα κοινωνικά δίκτυα συνεχώς μεγαλώνουν και η επιρροή των χρηστών (κόμβων) που συμμετέχουν σε αυτά αυξάνεται, είναι ανάγκη να έχουμε αξιόπιστους και δίκαιους τρόπους να αξιολογούμε τη σημαντικότητά τους,**

Στην διπλωματική αυτή θα μελετηθούν παραλλαγές της εργασίας [1]. Στην [1], όλοι οι κόμβοι ενός δικτύου είναι δίκαιοι, δηλαδή, κάνουν δίκαιους περιπάτους. Στη διπλωματική:

(α) θα εξετάσουμε την περίπτωση όπου μόνο συγκεκριμένα υποσύνολα κόμβων κάνουν δίκαιους τυχαίους περιπάτους, ή/και (β) τη αντικατάσταση του τυχαίου περιπάτου σε γνωστούς αλγορίθμους embedding με το δίκαιο τυχαίο περίπατο.

Θα χρειαστεί:

- (1) Θεωρητική κατανόηση του PageRank
- (2) Υλοποίηση της παραλλαγής – υπάρχει ήδη υλοποίηση του βασικού δίκαιου περιπάτου (σε C και Python), και
- (3) Πειραματική αξιολόγηση με χρήση συνθετικών και πραγματικών δικτύων (αυτά τα δίκτυα τα έχουμε ήδη δημιουργήσει/συλλέξει). Τα δίκτυα που θα χρησιμοποιηθούν

αφορούν πραγματικά κοινωνικά δίκτυα και οι μετρήσεις θα δείξουν κατά πόσο είναι δίκαια στην παρούσα μορφή τους και κατά πόσο ο νέος δίκαιος αλγόριθμος τα βελτιώνει.

Αναφορές

[1] Sotiris Tsioutsoulis, Evaggelia Pitoura, Panayiotis Tsaparas, Ilias Kleftakis, Nikos Mamoulis: Fairness-Aware PageRank. WWW 2021: 3815-3826

[2] Περιγραφή του PageRank υπάρχει σε πολλά βιβλία, αλλά και στο Wikipedia. Δείτε το Κεφάλαιο 5 του βιβλίου: Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman: Mining of Massive Datasets, 2nd Ed. Cambridge University Press 2014
<http://infolab.stanford.edu/~ullman/mmds/ch5.pdf>

Διπλωματική 2: Πειραματική σύγκριση embedding για temporal graphs

Τα graph embeddings χρησιμοποιούν μηχανική μάθηση για να απεικονίσουν τους κόμβους ενός δικτύου σε πολύ-διάστατα διανύσματα. Η βασική ιδέα αυτών των απεικονίσεων είναι παρόμοιοι κόμβοι (πχ, κόμβοι με πολλούς κοινούς γείτονες) να απεικονίζονται σε παρόμοια διανύσματα.

Σε αυτή τη διπλωματική θα γίνει πειραματική αξιολόγηση τεχνικών embeddings που έχουν προταθεί για χρονικά μεταβαλλόμενα γραφήματα. Θα αξιολογηθούν 5-10 τέτοιοι αλγόριθμοι σε 10 σύνολα δεδομένων.

Η αξιολόγηση θα αφορά στην απόδοση των τεχνικών στα παρακάτω προβλήματα: (1) χρονικές συστάσεις (temporal recommendations), (2) πρόβλεψη συνδεσιμότητας (reachability) και (3) εκτίμηση απόστασης (shortest path distance).

Διπλωματική 3: Πρόβλεψη αλλαγών σε ομάδες (communities) σε κοινωνικά δίκτυα

Σε πολλά δίκτυα παρατηρείται η δημιουργία ομάδων. Καθώς τα δίκτυα αλλάζουν στο χρόνο μεταβάλλεται και η συμμετοχή των κόμβων σε αυτές τις ομάδες. Σε αυτή τη διπλωματική θα επεκταθεί μια μέθοδος που έχουμε προτείνει στο [1] για την πρόβλεψη της μετακίνησης των κόμβων.

Αναφορές

[1] Evangelia Tsoukanara, Georgia Koloniaris, Evaggelia Pitoura, Should I Stay or Should I Go: Predicting Changes in Cluster Membership (under review)