

## Πέμπτη Σειρά Ασκήσεων

Η άσκηση πρέπει να παραδοθεί μέχρι την **Δευτέρα 11 Φεβρουαρίου**. Μπορείτε να πάρετε κάποιες επιπλέον μέρες παράταση αν έχετε εξετάσεις πριν από την ημερομηνία παράδοσης. Για καθυστερημένες ασκήσεις ισχύει η πολιτική στη σελίδα του μαθήματος. Λεπτομέρειες για το turn-in στη σελίδα του μαθήματος.

### Άσκηση 1

Αποδείξτε ότι για ένα μη κατευθυνόμενο γράφο η κατανομή σύγκλισης (stationary distribution) ενός τυχαίου περιπάτου είναι ανάλογη του βαθμού του κάθε κόμβου. Δηλαδή αν  $P$  είναι ο πίνακας μετάβασης του τυχαίου περιπάτου, και  $\pi$  η κατανομή σύγκλισης για την οποία ισχύει ότι  $\pi = \pi \cdot P$ , δείξτε ότι για τον κόμβο  $i$ , η πιθανότητα  $\pi_i$  είναι ανάλογη του  $d_i$ , όπου  $d_i$  είναι ο αριθμός των ακμών με άκρο την κορυφή  $i$ .

### Άσκηση 2

Για την άσκηση αυτή θα πρέπει να κάνετε τις ασκήσεις 4 και 6 από το κεφάλαιο 14 του βιβλίου “Networks, Crowds and Markets” των D. Easley και J. Kleinberg. Το link στο βιβλίο είναι στη σελίδα του μαθήματος. Η εκφώνηση των ασκήσεων είναι στα αγγλικά. Επικοινωνήστε μαζί μου αν έχετε απορίες.

### Άσκηση 3

Στην άσκηση αυτή θα πειραματιστείτε με τους αλγόριθμους PageRank και HITS. Για την υλοποίηση του PageRank, θα έχετε ομοιόμορφο jump vector, και πιθανότητα επανεκκίνησης 0.15. Εφαρμόστε τους αλγόριθμους σε ένα κατευθυνόμενο γράφο από το Stanford Network Analysis Project (η διεύθυνση είναι στη σελίδα του μαθήματος). Προτείνεται ο γράφος wiki-vote αλλά μπορείτε να χρησιμοποιήσετε όποιον γράφο θέλετε. Μπορείτε επίσης να προτείνετε κάποιον άλλο γράφο που σας ενδιαφέρει. Η υλοποίηση μπορεί να είναι σε οποιαδήποτε γλώσσα. Παραδώσετε τον κώδικα σας, αρχεία με τους κόμβους του δικτύου ταξινομημένους με βάση τον κάθε αλγόριθμο και την τιμή για κάθε κόμβο (το authority value στην περίπτωση του HITS), και μια σύντομη αναφορά που θα εξηγήει πώς δουλεύει ο κώδικας σας. Επίσης, συγκρίνετε τα πρώτα 20 αποτελέσματα του HITS και του PageRank.

## Άσκηση 4

Ας υποθέσουμε ότι έχετε ένα supermarket, και για κάθε προϊόν του καταστήματος έχετε τη λίστα των πελατών που έχουν αγοράσει το προϊόν. Θέλετε να δημιουργήσετε ένα κατάλογο για να στείλετε στους πελάτες του καταστήματος με προσφορές για  $K$  προϊόντα. Αν ο κατάλογος περιέχει ένα από τα προϊόντα που έχει αγοράσει ο πελάτης στο παρελθόν τότε ο πελάτης είναι *ικανοποιημένος* (και θα ξανάρθει στο κατάστημα). Θέλετε να διαλέξετε τα  $K$  προϊόντα του καταλόγου ώστε να μεγιστοποιήσετε τον αριθμό των πελατών που θα είναι ικανοποιημένοι. Δείξτε ότι το πρόβλημα είναι NP-hard, και ότι υπάρχει προσεγγιστικός αλγόριθμος με λόγο προσέγγισης  $(1-1/e)$ .

## Άσκηση 5 (Bonus)

Στην άσκηση αυτή θα πειραματιστείτε με τυχαίους περιπάτους με απορροφητικούς κόμβους. Από τη σελίδα του μαθήματος μπορείτε να κατεβάσετε το *movie-actor graph*, ένα διμερή γράφημα με κορυφές για ταινίες (2000-2010) και για ηθοποιούς και ακμές μεταξύ τους αν κάποιος ηθοποιός έχει παίξει σε μία ταινία. Διαλέξτε πέντε ή παραπάνω ταινίες δράσης, και πέντε ή παραπάνω κωμωδίες, και κάνετε τους αντίστοιχους κόμβους απορροφητικούς. Δώστε τιμή 1 στις κωμωδίες και τιμή -1 στις ταινίες δράσης. Χρησιμοποιώντας τη διαδικασία που περιγράψαμε στο μάθημα υπολογίστε την αναμενόμενη τιμή για όλους τους υπόλοιπους κόμβους (ταινίες και ηθοποιούς). Θα παραδώσετε το κώδικα σας και μια μικρή αναφορά που εξηγεί τον κώδικα. Επίσης ένα αρχείο με τους κόμβους που έχετε αρχικοποιήσει ως θετικούς και αρνητικούς, και δύο αρχεία με τις ταινίες και τους ηθοποιούς και τις τιμές που υπολογίσατε. Γράψετε επίσης και μια σύντομη αναφορά με τις παρατηρήσεις σας ως προς τα αποτελέσματα του προγράμματος σας.