

3ο Σύνολο Ασκήσεων

Ημερομηνία Παράδοσης: 27 Ιουνίου 2008 (μέχρι τις 13μμ στη βοήθό του μαθήματος)
Ενότητα: PageRank, HITS, Αποθήκες Δεδομένων

Ποσοστό επί του τελικού βαθμού: **25 %** για όσους ασχοληθούν με τις ασκήσεις (*)
12 % για τους υπόλοιπους

Οι ασκήσεις με χαρακτηρισμό **A** είναι ατομικές, ενώ οι ασκήσεις με χαρακτηρισμό **Δ** μπορεί να γίνουν σε ομάδες έως 2 ατόμων.

Οι ασκήσεις με **(*)** είναι προαιρετικές με την παρακάτω έννοια: μπορείτε να τις παραδώσετε αντί τελικής εξέτασης. Για αυτούς που θα επιλέξουν να ολοκληρώσουν όλες τις ασκήσεις (δηλαδή και τις προαιρετικές), ο τελικός βαθμός τους θα προκύψει από τον βαθμό τους στα σύνολα ασκήσεων. Για τους υπόλοιπους, οι ασκήσεις θα συμμετέχουν με ποσοστό 50% στον τελικό τους βαθμό.

Άσκηση 1 [Δ, ()]

Για τα (α) και (β) μπορείτε να χρησιμοποιήσετε το απλό PageRank (δηλαδή, χωρίς τον παράγοντα απόσβεσης).

(α) Δώστε ένα παράδειγμα γράφου που ο PageRank και ο HITS δίνουν το ίδιο αποτέλεσμα για κάποιο θέμα (δηλαδή, η σελίδα (κόμβος) που έχει το μεγαλύτερο PageRank έχει και το μεγαλύτερο authority).

(β) Δώστε ένα παράδειγμα γράφου που ο PageRank και ο HITS δε δίνουν το ίδιο αποτέλεσμα για κάποιο θέμα (δηλαδή, η σελίδα που έχει το μεγαλύτερο PageRank δεν έχει και το μεγαλύτερο authority).

(γ) Χρησιμοποιήστε ένα από τα πολλά εργαλεία που υπάρχουν για τον υπολογισμό του PageRank και δώστε το PageRank της σελίδας του Τμήματος μας, του Πανεπιστημίου, του google και - αν έχετε προσωπική σελίδα - της σελίδας σας. Η απάντησή σας να περιλαμβάνει απλώς το όνομα του εργαλείου που χρησιμοποιήσατε και τους 3 (4) αριθμούς.

Άσκηση 2 [A, (*)]

(α) Εξηγήστε το ρόλο του παράγοντα απόσβεσης (damping factor - d).

(β) Θεωρείστε έναν γράφο (όποιον θέλετε) με 5 κόμβους και δώστε το PageRank των κόμβων του με και χωρίς τη χρήση του παράγοντα απόσβεσης. Θεωρείστε $d = 0.85$.

Άσκηση 3 [Δ, ()]

Θεωρείστε μια αποθήκη δεδομένων για ένα Πανεπιστήμιο που αποτελείται από τις εξής τέσσερις διαστάσεις: Φοιτητής, Μάθημα, Εξάμηνο, Καθηγητής και δύο μέτρα (measures) Αριθμός_Βαθμών και Μέσος_Βαθμός. Για το χαμηλότερο εννοιολογικό επίπεδο (δηλαδή, για ένα συγκεκριμένο φοιτητή, μάθημα, εξάμηνο και καθηγητή), ο Μέσος_Βαθμός είναι ο βαθμός που πήρε ο φοιτητής και ο Αριθμός_Βαθμών είναι ίσος με 1. Σε υψηλότερα επίπεδα, ο Μέσος_Βαθμός είναι ο μέσος όρος των βαθμών που αντιστοιχεί στο συγκεκριμένο συνδυασμό και αντίστοιχα ο Αριθμός_Βαθμών ο αριθμός μαθημάτων που αντιστοιχεί στο συγκεκριμένο συνδυασμό.

(α) Δώστε μια πιθανή εννοιολογική ιεραρχία για κάθε μία από τις διαστάσεις.

(β) Σχεδιάστε ένα σχήμα Αστέρι (star schema) για την αποθήκη.

(γ) (προαιρετικά + 10) Σχεδιάστε ένα σχήμα Νιφάδας (snowflake) για την αποθήκη.

(δ) Δώστε τις βασικές λειτουργίες (αλγεβρικές πράξεις) που χρειάζονται για να πάρουμε το μέσο όρο των βαθμών όλων των φοιτητών στα μαθήματα του Τμήματος Πληροφορικής ξεκινώντας από το βασικό κυβειδές.