Ιωάννης Γιαννακός 4970

Ερώτημα 1 - Απόδειξη (ερώτημα 3)

Πρέπει να δείξω ότι το διάνυσμα P_u είναι ο μέσος όρος των $P_1, P_2, ..., P_n$ δηλαδή $P_u = \frac{1}{n}\sum_{i=1}^{n} P_i$, χρησιμοποιώντας $P_v^T = v^T Q$

Αρχικά, γνωρίζω ότι $P_u \rightarrow$ Pagerank διάνυσμα με ομοιόμορφο v^T
$P_i \rightarrow$ Personalized Pagerank διάνυσμα
και $v^T \rightarrow$ jump vector, το οποίο στο P_u διάνυσμα έχει την

μορφή $v^T = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$, εφόσον είναι ομοιόμορφο, και

στο P_i διάνυσμα έχει την μορφή $v_i^T = (0, 0, ..., 0, 1, ..., 0)$ με το 1 στην i-θέση, εφόσον γνωρίζουμε ότι το jump vector σε αυτή την περίπτωση δίνει όλη την πιθανότητα στον κόμβο i.

Στη συνέχεια, αν στην δοθείσα σχέση $P_v^T = v^T Q$ αντικαταστήσω όπου P_v^T το P_i^T, δηλαδή το personalized Pagerank διάνυσμα

Θα έχω: $P_i^T = v_i^T Q$ (1), όπου το v_i^T είναι το jump vector του personalized Pagerank διανύσματος, το οποίο αφού δίνει όλη τη πιθανότητα στον κόμβο i, θα είναι ένα διάνυσμα με μηδενικά παντού, εκτός από την θέση i που θα έχει την τιμή 1.

Συνεπώς από την σχέση (1) το P_i^T θα είναι η i-οστή γραμμή του πίνακα Q.

Επομένως, το Pagerank διάνυσμα με ομοιόμορφο jump vector P_u είναι ο μέσος όρος των personalized Pagerank διανυσμάτων P_i για όλους τους κόμβους i.

$$\text{Δηλαδή } P_u = \frac{1}{n}\sum_{i=1}^{n} P_i$$

1) $P_v^T = (1-a)P_v^T \cdot P + av^T \quad \rightarrow$ φόρμουλα για πίνακα Q

Αρχικά, γνωρίζουμε ότι $P_v^T = v^T \cdot Q$, άρα η παραπάνω σχέση γράφεται

$v^T Q = (1-a)v^T Q \cdot P + av^T \implies v^T Q - (1-a)v^T Q \cdot P = av^T$

$\implies Q - (1-a)Q \cdot P = a \implies Q[1-(1-a)\cdot P] = a \implies \boxed{Q = \dfrac{a}{1-(1-a)\cdot P}}$

2) $P_v^T = v^T Q \rightarrow$ τι ισχύει για τις γραμμές του πίνακα Q ;
(ως προς τα personalized Pagerank vectors)

Αν η παραπάνω σχέση εφαρμοστεί για κάθε γραμμή i του πίνακα Q, τότε θα έχουμε $P_{v_i}^T = v_i^T Q$, όπου P_{v_i} είναι το personalized Pagerank διάνυσμα με jump vector που δίνει όλη τη πιθανότητα στον κόμβο i (v_i)

Συνεπώς κάθε γραμμή του Q αντιστοιχεί στο personalized Pagerank διάνυσμα για τον αντίστοιχο κόμβο, δηλαδή οι γραμμές του Q αντιπροσωπεύουν τα personalized Pagerank διανύσματα για κάθε κόμβο.

4) Σε οποιοδήποτε jump vector v (όχι απαραίτητα ομοιόμορφο), πως εκφράζουμε το P_v συναρτήσει των P_i ;

Γενικά, το Pagerank διάνυσμα με οποιοδήποτε jump vector v μπορεί να εκφραστεί ως γραμμικός συνδυασμός των personalized Pagerank διανυσμάτων P_i με βάση τον πίνακα Q που έχει καθοριστεί από τον jump vector v, μέσω της σχέσης $P_v^T = v^T \cdot Q$, η οποία μας λέει ότι κάθε στοιχείο P_{v_i} του Pagerank διανύσματος είναι το αποτέλεσμα του γινομένου της γραμμής i του πίνακα Q με το αντίστοιχο personalized Pagerank διάνυσμα P_i