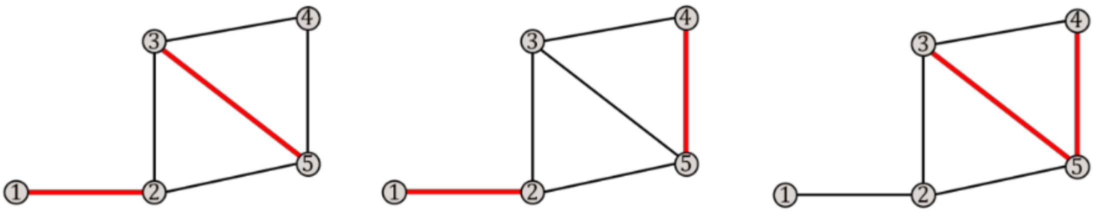
Set 3

Άσκηση 1 (Ταιριάσματα)

Έστω $G = (V, E)$ ένα απλό γράφημα. Ένα σύνολο ακμών $M \subseteq E$ του G αποτελεί ένα ***ταίριασμα*** αν στο υπογράφημα $G_M = (V, M)$ του G κάθε κόμβος έχει βαθμό ίσο με 0 ή 1. (Δηλαδή κάθε κόμβος $v \in V$ έχει το πολύ ένα γείτονα στο G_M.) Το παρακάτω σχήμα απεικονίζει δύο ταιριάσματα M_1 και M_2 ενός γραφήματος, καθώς και τις ακμές της συμμετρικής διαφοράς τους $N = M_1 \oplus M_2$.



$M_1 = \{\{1,2\}, \{3,5\}\}$ $\qquad$ $M_2 = \{\{1,2\}, \{4,5\}\}$ $\qquad$ $N = (M_1 - M_2) \cup (M_2 - M_1)$

Το ταίριασμα M είναι *τέλειο* αν κάθε κορυφή έχει βαθμό στο G_M ακριβώς 1.

α. Έστω δύο διαφορετικά ταιριάσματα M_1 και M_2 του G, και έστω $N = M_1 \oplus M_2 = (M_1 - M_2) \cup (M_2 - M_1)$ η συμμετρική διαφορά τους. Δηλαδή, το N είναι το σύνολο των ακμών που είτε ανήκουν στο M_1 αλλά όχι στο M_2, είτε ανήκουν στο M_2 αλλά όχι στο M_1. Δείξτε ότι το $G_N = (V, N)$ είναι υπογράφημα του G, στο οποίο κάθε συνεκτική συνιστώσα είναι μονοπάτι ή απλός κύκλος άρτιου μήκους.

Υπόδειξη: Δείξτε ότι κάθε κορυφή του G_N έχει βαθμό το πολύ ίσο με 2.

β. Χρησιμοποιήστε το παραπάνω ερώτημα για να δείξετε ότι το αν το G είναι δένδρο (δηλαδή συνεκτικό γράφημα χωρίς κύκλους), τότε έχει το πολύ ένα τέλειο ταίριασμα.

γ. Σχεδιάστε ένα αποδοτικό αλγόριθμο που να υπολογίζει ένα μέγιστο ταίριασμα σε ένα δένδρο G με n κόμβους. Ποιος είναι ο χρόνος εκτέλεσης του αλγορίθμου σας;

Άσκηση 2 (Αναγωγές στον υπολογισμό μέγιστου ταιριάσματος)

α. Μια ***κάλυψη ακμών*** ενός συνεκτικού γραφήματος $G = (V, E)$ είναι ένα υποσύνολο $R \subseteq E$ των ακμών του G, τέτοιο ώστε για κάθε κόμβο $v \in V$, να υπάρχει τουλάχιστον μια ακμή $\{v, w\}$ στο R (όπου w κάποιος γείτονας του v στο G). Περιγράψτε ένα αποδοτικό αλγόριθμο (πολυωνυμικού χρόνου) για τον υπολογισμό μιας *ελάχιστης κάλυψης ακμών*, δηλαδή μιας κάλυψης ακμών με ελάχιστο μέγεθος.

Υπόδειξη: Δείξτε πρώτα ότι το μέγεθος μιας ελάχιστης κάλυψης ακμών R^* είναι $|R^*| = |V| - |M^*|$, όπου M^* ένα μέγιστο ταίριασμα στο G.

β. Μας δίνεται ένα διμερές γράφημα $G = (V \cup W, E)$ και ένας θετικός ακέραιος b_v για κάθε κορυφή $v \in V \cup W$. Ένα ***b-ταίριασμα*** είναι ένα υποσύνολο $M \subseteq E$ ακμών, τέτοιο ώστε κάθε κορυφή v να προσπίπτει σε το πολύ b_v ακμές του M. (Το τυπικό πρόβλημα διμερούς ταιριάσματος αντιστοιχεί στην περίπτωση όπου $b_v = 1$ για κάθε $v \in V \cup W$.) Αποδείξτε ότι το πρόβλημα του υπολογισμού ενός μέγιστου b-ταιριάσματος στο G ανάγεται στο πρόβλημα του υπολογισμού ενός (τυπικού) μέγιστου διμερούς ταιριάσματος σε ένα μεγαλύτερο γράφημα. Η αναγωγή σας θα πρέπει να τρέχει σε πολυωνυμικό χρόνο στο μέγεθος του G και στην τιμή $b^* = \max\{b_v : v \in V \cup W\}$.

Άσκηση 3 (Πιθανοτική μέθοδος)

Έστω $G = (V, E)$ ένα συνεκτικό γράφημα με n κόμβους και m ακμές. Ένα *κυρίαρχο σύνολο* του G είναι ένα υποσύνολο $S \subseteq V$ των κόμβων του γραφήματος, τέτοιο ώστε κάθε κόμβος του G είτε ανήκει στο S είτε έχει ένα γειτονικό κόμβο στο S. Έστω ότι κάθε κόμβος του G έχει βαθμό τουλάχιστον δ. Δείξτε ότι το G έχει ένα κυρίαρχο σύνολο με το πολύ $\frac{n(1+\lg(\delta+1))}{\delta+1}$ κόμβους.

Υπόδειξη: Θεωρήστε την ακόλουθη πιθανοτική κατασκευή ενός κυριάρχου συνόλου S. Κατασκευάζουμε ένα σύνολο κόμβων T επιλέγοντας κάθε κόμβο του G με πιθανότητα p. Έστω A το σύνολο των άτυχων κόμβων που δεν ανήκουν στο T και δεν έχουν γείτονα στο T. Θέτουμε $S = T \cup A$. Ποιο είναι το αναμενόμενο μέγεθος του S; Δείξτε ότι η πιθανότητα που ελαχιστοποιεί το $E[S]$ είναι $p = \frac{lg(\delta+1)}{\delta+1}$. Στην ανάλυσή σας μπορείτε να χρησιμοποιήσετε την ανισότητα $1 - x \leq e^{-x}$.

Άσκηση 4 (Τυχαιοκρατικοί αλγόριθμοι)

Η κεντρικότητα ενός κόμβου v σε ένα συνεκτικό γράφημα $G = (V, E)$ με n κόμβους και m ακμές ορίζεται ως

$$c_v = \frac{n-1}{\sum_{u \in V} d(u, v)}$$

όπου $d(u, v)$ η απόσταση μεταξύ των κόμβων u και v στο G.

α. Περιγράψτε ένα απλοϊκό αλγόριθμο υπολογισμού της κεντρικότητας όλων των κόμβων του G. Ποιος είναι ο χρόνος εκτέλεσης του;

β. Ένας πιθανοτικός αλγόριθμος υπολογισμού της κεντρικότητας κάθε κόμβου βασίζεται στη τυχαία δειγματοληψία. Έστω k μια παράμετρος, η οποία καθορίζει την πιθανότητα επιτυχίας του αλγόριθμου. Ο αλγόριθμος εκτελεί τα παρακάτω βήματα:

> **for** $i = 1$ **to** k
>
> επίλεξε ένα κόμβο $u_i \in V$ ομοιόμορφα τυχαία
>
> υπολόγισε τις αποστάσεις $d(u_i, v)$ με αφετηρία τον κόμβο u_i
>
> **end**
>
> **for** $v \in V$
>
> υπολόγισε την εκτίμηση της κεντρικότητας του v ως
>
> $$\hat{c}_v = \frac{k(n-1)}{n \sum_{i=1}^{k} d(u_i, v)}$$
>
> **end**

Δείξτε ότι η αναμενόμενη τιμή της ποσότητας $1/\hat{c}_v$ είναι $1/c_v$.

Υπόδειξη: Θεωρήστε πρώτα την περίπτωση $k = 1$.

γ. Μπορούμε να δώσουμε ένα καλό φράγμα για τη πιθανότητα ο παραπάνω αλγόριθμος να μην υπολογίσει καλές εκτιμήσεις της κεντρικότητας κάθε κόμβου χρησιμοποιώντας το παρακάτω φράγμα του Hoeffding. Έστω ανεξάρτητες τυχαίες μεταβλητές $x_1, x_2, \dots, x_k$, όπου $a_i \le x_i \le b_i$, με αναμενόμενη μέση τιμή $\mu = \mathbf{E}\left[\sum_{i=1}^{k} x_i / k\right]$.

Τότε, για $\xi > 0$ ισχύει

$$\Pr\left[\left|\frac{\sum_{i=1}^{k} x_i}{k} - \mu\right| \ge \xi\right] \le 2\exp\left(\frac{-2k^2\xi^2}{\sum_{i=1}^{k}(a_i - b_i)^2}\right)$$

όπου $\exp(x) = e^x$.

Στην περίπτωση μας, θέτουμε $x_i = \frac{n\, d(u_i, v)}{n-1}$, οπότε έχουμε $a_i = 0$, $b_i = \frac{n\Delta}{n-1}$, όπου Δ η διάμετρος του γραφήματος (δηλαδή η μέγιστη απόσταση δύο κόμβων) και $\mu = \mathbf{E}\left[\sum_{i=1}^{k} x_i / k\right] = 1/c_v$.

Με βάση το παραπάνω φράγμα, δείξτε ότι για $k \approx \frac{\ln n}{\varepsilon^2}$ ισχύει

$$\Pr\left[\left|\frac{1}{\hat{c}_v} - \frac{1}{c_v}\right| \ge \varepsilon\Delta\right] \le \frac{1}{n^2}.$$

Στη συνέχεια δείξτε ότι η πιθανότητα να υπάρχει κάποιος κόμβος v για τον οποίο $\left|\frac{1}{\hat{c}_v} - \frac{1}{c_v}\right| \ge \varepsilon\Delta$ είναι το πολύ $1/n$.

Άσκ 1 (α) • $N = M_1 \oplus M_2 = \underbrace{(M_1 - M_2)}_{\text{Μόνο ακμές του υποσυνόλου } M_1} \cup \underbrace{(M_2 - M_1)}_{\text{Μόνο ακμές του υποσυνόλου } M_2}$ ή

(Βλάχος Δημήτριος)

> Συνεπώς ∀ ακμή του συνόλου $N \in E(G)$, είναι ∀ κόμβος του G θα είναι μέλος το πολύ μιας ακμής του ταιριάσματος M_1 ή θα είναι μέλος το πολύ μιας ακμής του ταιριάσματος M_2 (∀ ακμή αποτελείται από 2 κόμβους. Εξ ορισμού).

> Το G_N είναι subgraph του G.

• Τώρα πρέπει να δείξουμε ότι ∀ connected component είναι path ή simple cycle with even length.

> ∀ node στα σύνολα M_1, M_2 (δηλ. οι node που προκύπτουν στις ακμές των 2 συνόλων) έχει βαθμό $=0$ ή βαθμό $=1$ από κάποιον-ορισμό του ταιριάσματος. Δηλ. έστω $u \in V(G)$ χωρίς βλάβη της γενικότητας.

> Αν ο u δεν είναι μέλος κάποιας ακμής του συνόλου N τότε $deg(u)$ στο $G_N = 0$

> Αν είναι μέλος μιας ακμής του συνόλου N τότε $deg(u)$ στο $G_N = 1$

> Επειδή ο u μπορεί να έχει τη μια προσκείμενη ακμή στο M_1 & την άλλη προσκείμενη ακμή στο M_2 $\Longrightarrow$
∀ κορυφή V_N του $G_N(V_N, E_N)$ θα έχει $degree(u) \leq 2$

> Συνεπώς δείξαμε ότι ο βαθμός ∀ κόμβου στο $G_N \leq 2$.
Άρα οι connected components του G_N είναι μονοπάτια ή κύκλοι.

> Μένει να δείξουμε ότι ο κύκλος είναι even length.

> ∀ κύκλος προκύπτει από τον συνδυασμό ακμών που υπάρχουν στα 2 ταιριάσματα (M_1 & M_2) εναλλάξ.
Pattern: 1 ακμή από το 1ο ταίριασμα, 1 ακμή από το 2ο ταίριασμα, ...

> $\underbrace{num1}_{\text{ακμές από } M_1} + \underbrace{num2}_{\text{ακμές από } M_2} = even\text{-}num$, $num1, num2 \in \mathbb{N}$.

> Άρα άρτιου μήκους κύκλος.

$\boxed{\text{Ασκ1}}$ (β) • Δέντρο ορισμός:

> ∄ κύκλος
> ∀ συνεκτικό υπογράφημα δεν έχει κύκλο (αφού είναι tree)

• Έστω ότι το G ∃ ≥ 2 τέλεια Matchings.

> Έστω M_1 & M_2 είναι τέλεια ταιριάσματα
> Τότε ο γράφος θα έχει ≥ 2 μονοπάτια για τη σύνδεση από 1 σημείο u σε 1 σημείο v $u, v \in V(G)$ $u \neq v$
> Όμως τότε ο γράφος δεν θα είναι Tree
> ⇒ Άτοπο: Δεν γίνεται να έχουμε παραπάνω από 1 ταίριασμα γιατί θα καταστρατηγούσαμε τον ορισμό των δέντρων.
> ⇒ $\boxed{\text{If } G \text{ is a tree} \Rightarrow \leq 1 \text{ Τέλειο Ταίριασμα}}$

> Δηλ: Θα υπήρχε επικάλυψη μεταξύ κόμβων στα ταιριάσματα. Κάτι που προφανώς δεν επιτρέπεται. (∀ κόμβος σε ένα τέλειο ταίριασμα θα πρέπει να ταιριάζει ακριβώς 1 φορά)

Απάντηση (γ)

```
Max_Matching_Tree {
    Input: G(V,E) , n
    Output: Μέγιστο_Ταίριασμα() // list

    n = |V|;
    Μέγιστο_Ταίριασμα = new(); // list initialization
    visited = [False, False, ... , False]; // size n

    for(∀ u in V(G)){
        if (visited[u] == False){
            DFS(u);     // DFS usage
        }
    }
    return (Μέγιστο_Ταίριασμα);
}

DFS(u) {
    visited[u] = True;
    for(v in G[u]){ // G = Tree ⟺ unique path from v
        if (visited[v] == False){
            DFS(v); // recursive usage
            if((v,u) ∉ Μέγιστο_Ταίριασμα){
                Μέγιστο_Ταίριασμα.add((v,u));
            }
        }
    }
}
```

DFS complexity $O(|V| + |E|)$

Θα επισκεφτεί ∀ κόμβο & ∀ ακμή ακριβώς μία φορά κατά την Depth First Search διερεύνηση $\Longrightarrow$ Γραμμικός

Ασκ3 • Θα ακολουθήσουμε την υπόδειξη.

> Χτίζουμε ένα σύνολο κόμβων T (υποσύνολο του $V(G)$):
> Επιλέγουμε τους κόμβους $\in V(G)$ πιθανοτικά & ανεξάρτητα τον έναν από τον άλλον.
> Ίδια πιθανότητα επιλογής $\forall$ κόμβου $= p$, άγνωστη προς το παρόν.

> Για το σύνολο A των ατυχών κόμβων:
> $\forall$ κόμβο $\in A$ — $\notin$ στον T & δεν έχει γειτονικό κόμβο $\in T$. } συνθήκες που ισχύουν ταυτόχρονα
> Συνεπώς κανένα στοιχείο του συνόλου A δεν ανήκει στο κυρίαρχο σύνολο

> Ορίζουμε τ.μ.: "X = # κόμβων που $\in T$."

$$X_i = \begin{cases} 0, & \text{ο i-οστός κόμβος } \in T \\ 1, & \text{ο i-οστός κόμβος } \notin T \end{cases}$$

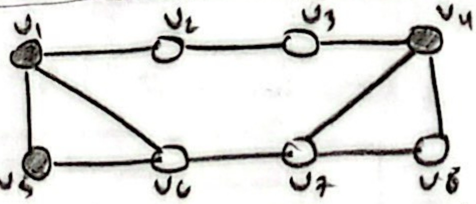
$$X = \sum_{i=1}^{n} X_i$$

> $E[X] = E[|T|]$, $|T|$: #στοιχείων του συνόλου

$$E[|T|] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Pr[X_i = 1] = \sum_{i=1}^{n} p = n \cdot p$$

Από λοιπόν είναι το αναμενόμενο $|T|$ που εξαρτάται από την p.

> ΙΙΧ



} Έστω $T = \{v_1, v_4, v_5\}$
v_2, v_6 γειτονικά του v_1
v_3, v_7, v_8 γειτονικά του v_4
Άρα $A = \varnothing$.

Συνεπώς: $T \cup A \subseteq V(G)$ στο παράδειγμά μας.

> ΙΙΧ2

Στο παράδειγμα όπου $\forall v_i$ έχει $degree(v_i) = 2$ εκτός των 2 ακραίων που έχουν $degree = 1$
Όποια επιλογή κόμβων να γίνει για το σύνολο $T \Rightarrow$
$T \cup A \subseteq V(G)$

> Συνεπώς μπορούμε χωρίς αυστηρή απόδειξη, διαισθητικά να καταλάβουμε ότι $T \cup A \subseteq V(G)$ πάντα.
($|S| = |T \cup A|$)

Ασκ 3 συνέχεια

> Αυτό μας "δείχνει" κατά κάποιο τρόπο ότι η πιθανότητα ενός κόμβου $u \in A$ εξαρτάται άμεσα από το degree του.

> $E[|A|]$: όμοια είναι το αναμενόμενο # κορυφών u στο A.

> Πρέπει $u \notin T \Rightarrow Pr = \underline{1-p}$
> Πρέπει να μην επιλέγει κανένας γείτονας ($\exists$ τουλ δ) στο T

Οι δύο παραπάνω συνθήκες ισχύουν για 1 κόμβο, έστω u.

> Η πιθανότητα ότι κανένας γείτονας από τους $d \geq \delta$ είναι $Pr = (1-p)^d$

> $Pr\{u \in A\} = (1-p)(1-p)^d = (1-p)^{d+1} \leq (1-p)^{\delta+1}$

> $E[|A|] = n \cdot (1-p)^{\delta+1}$

> Άρα $E[|S|] = E[|T|] + E[|A|] = \boxed{n \cdot p + n \cdot (1-p)^{\delta+1}}$

Έστω: $f(p) = n \cdot p + n(1-p)^{\delta+1} = n(p + (1-p)^{\delta+1})$ (1)

$f'(p) = n + n(\delta+1) \cdot (1-p)^\delta (1-p)' =$

$= n + n(\delta+1)(1-p)^\delta(-1) \Longleftrightarrow$

$f'(p) = n \cdot \left(1 - (\delta+1)(1-p)^\delta\right)$

$f'(p) = 0 \Leftrightarrow 1 - (\delta+1)(1-p)^\delta = 0 \Leftrightarrow (\delta+1)(1-p)^\delta = 1 \Leftrightarrow$

$\Leftrightarrow (1-p)^\delta = \frac{1}{\delta+1} \Leftrightarrow 1-p = \frac{1}{(\delta+1)^{1/\delta}} \Longleftrightarrow$

$\boxed{p = 1 - \frac{1}{(\delta+1)^{1/\delta}}} = \boxed{1 - \left(\frac{1}{\delta+1}\right)^{1/\delta}}$

p		$1 - \frac{1}{(\delta+1)^{1/\delta}}$	
$f'(p)$	$-$		$+$
$f(p)$	$\searrow$		$\nearrow$

Η f παρουσιάζει ελάχιστο στο $p = 1 - \frac{1}{(\delta+1)^{1/\delta}}$

χρήση λογαρίθμου.

$\boxed{\text{Ασκ 3}}$ συνέχεια

> Τα πράγματα περιπλέκονται κάπως με χρήση λογαρίθμου. Αυτό που θέλουμε τώρα είναι να ελαχιστοποιήσουμε την $E[|S|]$ $(= F(p))$, δηλ. να βρούμε κατάλληλο p.

> Για $\delta \gg 0$: $\left(\frac{1}{\delta+1}\right)^{1/\delta} \approx e^{-\frac{1}{\delta}\log(\delta+1)} \Longrightarrow p \approx \boxed{\frac{\log(\delta+1)}{\delta+1}}$

$$E[|S|] = n \cdot \underbrace{\left[\frac{\log(\delta+1)}{\delta+1} + \left(1 - \frac{\log(\delta+1)}{\delta+1}\right)^{\delta+1}\right]}_{E[|T|] + E[|A|]} \xrightarrow{\frac{1}{e^x} \geq 1-x}$$

> $$\frac{1}{e^x} \geq 1-x \Longleftrightarrow e^{-x} \geq x \Longrightarrow$$
$$\left(1 - \frac{\log(\delta+1)}{(\delta+1)}\right)^{\delta+1} \leq e^{-\log(\delta+1)} = \frac{1}{\delta+1}$$

> Άρα: $E[|S|] \leq n\left(\frac{\log(\delta+1)}{(\delta+1)} + \frac{1}{\delta+1}\right) = \frac{n(1+\log(\delta+1))}{\delta+1}$

Συνεπώς θα έχουμε κυρίαρχο set S με αναμενόμενο μέγεθος το πολύ $\frac{n(1+\log(\delta+1))}{\delta+1}$ κόμβους

Άσκ 4 (a)

- $\sum_{u \in V} d(u,v) = \sum_{u \in V(G)} dist(u,v) = dist(v)$. (1)

- $\Longleftrightarrow$ "The sum of the distances between a specific node and all the nodes of a given graph $G=(V,E)$" (Ορισμός από Graph Theory Course)

- Έτσι για $\forall\, v \in V$ είναι εύκολο να υπολογίσουμε το dist(v) με χρήση κάποιου γνωστού αλγορίθμου. (Bellman-Ford, Dijkstra (για μη-αρνητικά βάρη) από n φορές ή Floyd-Warshall (πινακικές διαδρομές) 1 φορά)

- Η (1) προφανώς ισχύει για έμβαρα & για μη-έμβαρα γραφήματα.

- Αλγόριθμος:

"Prepare" for Dijkstra

```
Step(1): For(∀ vi ∈ V(G)) { // Initialize to a huge num.
              dist[vi] = ∞;  // array of size n
         }
                                    or an Algorithm that can
                                    do the same work
Step(2): // For ∀ node vi we Dijkstra & sum the
         // distances from vi to all other nodes
         For(∀ vi ∈ V(G)) {
           (2.1): Run Dijkstra |V| times to compute
                  d(u,vi) for all u's in V(G).
           (2.2): dist[vi] = Σ_{u∈V(G), u≠v} d(u,vi)
         }
Step(3):
         For(∀ vi ∈ V(G)) {
              C_vi = (n-1) / dist[vi];  // n = |V|
         }
```

Ασκ 4 (α) <u>συνέχεια</u>

Κόστος του Αλγορίθμου:

- Χρησιμοποιώντας τον Dijkstra με μία heap priority queue, ώστε να βρούμε την απόσταση μεταξύ 2 κόμβων:
  $O((n+m)*\log n)$
- Συνεπώς για το Step(2) όπου θα κάνουμε χρήση του Dijkstra n φορές: $\boxed{O(n*(n+m)*\log n)}$
- Step(1): Η πράξη μέσα στο loop έχει κόστος $O(1) \Longrightarrow$ Συνολικά έχουμε κόστος $\boxed{O(n)}$
- Step(3): Ακριβώς ίδια λογική με το Step(1).
  $O(1)$ για την πράξη (διαίρεση) μέσα στο loop.
  $\Longrightarrow$ Συνολικά για το step, έχουμε κόστος $O(1)$.
- Άρα ο Αλγόριθμος έχει κόστος:
  $\underline{O(n*(n+m)*\log n)}$
- <u>Σχόλιο</u>: Από το κόστος συνάγουμε το συμπέρασμα ότι για τεράστια γραφήματα με πολλές ακμές θα είναι δυστυχώς αργός ο Αλγόριθμός μας.
- <u>Σχόλιο 2</u>: Αν ο γράφος δεν έχει βάρη τότε μπορούμε να "δώσουμε" σε κάθε ακμή βάρος 1.
- <u>Σχόλιο 3</u>: Εναλλακτικά για undirected & unweighted graphs μπορούμε να κάνουμε χρήση του BFS.

Άσκ4 (β) • Έστω $k=1$ & έστω ένας κόμβος v_1.

- Υπολογίζουμε των $d(u,v_1)$ για $\forall\, u \in V(G)$ με χρήση του Αλγορίθμου που μας δόθηκε
- Θα κάνουμε την ακόλουθη αλλαγή μέσα στο 2ο loop:
  Έστω ότι υπολογίζουμε όπως είπαμε τη κεντρικότητα ενός κόμβου που τυχαία επιλέξαμε (ομοιόμορφη κατανομή)

$$\hat{C}_v = \frac{k\cdot(n-1)}{n\cdot d(u,v_1)} \overset{k=1}{=\!=} \frac{n-1}{n\, d(u,v_1)} \quad (\text{αυτό θέλουμε})$$

- Για $k=1$: Θέλουμε να δείξουμε ότι $\boxed{\frac{1}{\hat{C}_v} = \frac{1}{C_v}}$
- $\forall$ κόμβος επιλέγεται ισοπίθανα.
  Συνεπώς η αναμενόμενη τιμή της απόστασης $d(u,v) =$

$$\sum_{\forall u \in V(G)} \frac{d(u,v)}{n} = \frac{dist(v)}{n} \Longleftrightarrow E[d(u,v)] = \frac{1}{n}\sum_{\forall u \in V(G)} \frac{d(u,v)}{1}$$

- Η αντίστοιχη τιμή της $\hat{C}_v = \frac{k\cdot(n-1)}{dist(v)\cdot n}$
  Για $k=1$: $\Longrightarrow \frac{n-1}{dist(v)\cdot n}$ που ισοδυναμεί με τον υπολογισμό του C_v.

- Γενικεύουμε για $k>1$:
  > Υπολογίζουμε τις αποστάσεις πολλών τυχαίων κόμβων.
  > Έτσι θα έχουμε μια καλή τιμή για την κεντρικότητα $\hat{C}_v$ καθώς τώρα η $d(u,v)$ θα προσεγγιστεί από διάφορους κόμβους και όχι μόνο από τον v_1.

  > Χρήση της ιδιότητας: $E[\hat{C}_v] = C_v$ για $k \geq 1$
  $\Longleftrightarrow E\left[\frac{1}{\hat{C}_v}\right] = \frac{1}{C_v}$

Δηλ.:

> $C_v = \frac{n-1}{\sum_{u \in V(G)} d(u,v)}$ & $E\left[\frac{1}{\hat{C}_v}\right] = \frac{n}{n-1}\cdot\frac{1}{n}\sum_{u \in V(G)} d(u,v) = \frac{1}{n-1}\sum_{u \in V(G)} d(u,v)$

$\Longleftrightarrow \frac{1}{C_v} = \frac{\sum_{u \in V(G)} d(u,v)}{n-1}$ & $E\left[\frac{1}{\hat{C}_v}\right] = \frac{1}{C_v}$

Άσκ 4 (γ) > Έστω δοσμένες ανεξάρτητες τ.μ. $X_1, X_2, \dots, X_k$ με $x_i \in [a_i; b_i]$
& αναμενόμενη τ.μ. $\mu = E\left[\sum_{i=1}^{k} x_i/k\right]$, $a_i \le x_i \le b_i$

> $$\Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k} d(u,v_i) - E\left[\frac{1}{k}\sum_{i=1}^{k} x_i\right]\right| \ge \xi\right] \le 2\exp\left(\frac{-2k^2\xi^2}{\sum_{i=1}^{k}(a_i-b_i)^2}\right) \iff$$

$$\Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k} d(u,v_i) - \mu\right| \ge \xi\right] \le 2\exp\left(\frac{-2k^2\xi^2}{\sum_{i=1}^{k}(a_i-b_i)^2}\right)$$

> Για $a_i = 0$, $b_i = \frac{n\Delta}{n-1}$, $\mu = 1/c_v$

το θεώρημα του Hoeffding γίνεται:

$$\Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k} d(u,v_i) - \frac{1}{c_v}\right| \ge \xi\right] \le 2\exp\left(\frac{-2k^2\xi^2}{\sum_{i=1}^{k}\left(\frac{n\Delta}{n-1}\right)^2}\right)$$

Δ = Η μέγιστη απόσταση μεταξύ 2 κόμβων.

$$\sum_{i=1}^{k}\left(\frac{n\Delta}{n-1}\right)^2 = \left(\frac{n\Delta}{n-1}\right)^2 \sum_{i=1}^{k} 1 = k\cdot\left(\frac{n\cdot\Delta}{n-1}\right)^2 \approx \frac{\ln n}{\varepsilon^2}\left(\frac{n\cdot\Delta}{n-1}\right)^2$$

$$\iff \Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k} d(u,v_i) - \frac{1}{c_v}\right| \ge \xi\right] \le 2\exp\left(\frac{-2k^2\varepsilon^2}{k\left(\frac{n\Delta}{n-1}\right)^2}\right) =$$

$$= 2\exp\left(\frac{-2k\varepsilon^2}{\left(\frac{n\Delta}{n-1}\right)^2}\right) \xrightarrow{k = \frac{\ln n}{\varepsilon^2}}$$

$$\Pr\left[\left|\frac{1}{\hat{c}_v} - \frac{1}{c_v}\right| \ge \varepsilon\right] \le 2\exp\left(\frac{-2\left(\frac{\ln n}{\varepsilon^2}\right)^2\varepsilon^2}{\frac{\ln n}{\varepsilon^2}\left(\frac{n\Delta}{n-1}\right)^2}\right) \iff$$

$$\Pr\left[\left|\frac{1}{\hat{c}_v} - \frac{1}{c_v}\right| \ge \varepsilon\right] \le \frac{1}{n^2}$$

Άρα: $\Pr\left(\text{“} \exists \text{ κόμβος } v \text{ για τον οποίο ισχύει η } \left|\frac{1}{\hat{c}_v} - \frac{1}{c_v}\right| \ge 3\Delta \text{”}\right) \le \underbrace{1/n}_{n\cdot\frac{1}{n^2}}$

Άσκ2 (β) Βήματα Αναγωγής:

- Αντικαθιστούμε ∀ κορυφή v του αρχικού γραφήματος με b_v κορυφές v_i (Δηλ: $v_1, v_2, \dots, v_k$ για $b_v = k$)
- Συνδέουμε τις νέες κορυφές ως εξής:
  ∀ ακμή (u,v) του αρχικού γραφήματος θα αντικατασταθεί από όλες τις δυνατές ακμές: μεταξύ των κορυφών που αντικαθιστούν u ($u_1, u_2, \dots, u_k$ $b_u = k$), v ($v_1, v_2, \dots, v_l = b_v = l$).
- Στο γράφημα που προκύπτει ∀ κόμβος u: $\deg(u) = 1$.
  Άρα χρειαζόμαστε μια λύση για το μέγιστο διμερές ταίριασμα.

Input: Bipartite graph $G = (V \cup W, E)$, b_v για $\forall v \in V \cup W$.

> Δημιουργία του νέου γραφήματος $G' = (V' \cup W', E')$

> Για($\forall v \in V \cup W$){
Δημιούργησε b_v κόμβους που αντικαθιστούν το v στο G'
}

> Για($\forall$ ακμή $(u,v) \in E$) {
Δημιούργησε όλες τις πιθανές ακμές που σχηματίζονται από τους κόμβους που έχουν προηγουμένως αντικαταστήσει τους u, v κόμβους!
}

> Χρήση Αλγορίθμου για μέγιστο bipartite Matching στο G' (έστω M)

> Return M

(α) Ν.Δ.Ο. το μέγεθος μιας ελάχιστης κάλυψης ακμών R^* είναι: $|R^*| = |V| - |M^*|$

> Έστω M^* μέγιστο ταίριασμα στο G.

> Το πολύ 1 ακμή από το M^* ενώνει 2 κόμβους
$\Longleftrightarrow$ ∀ κόμβος καλύπτεται από το M^*.

> Οι ακμές του M^* δεν είναι δυνατό να καλύψουν περισσότερους κόμβους από το πλήθος των κόμβων που μπορούμε να καλύψουμε από 1 ταίριασμα.

> Με βάση γνωστό Θεώρημα: |μέγιστο ταίριασμα| = |ελάχιστη κάλυψη κόμβων σε διμερές γράφημα|