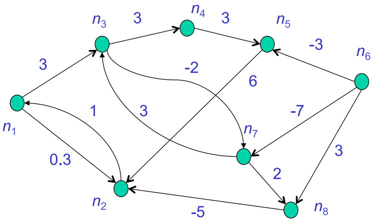
Εισαγωγή στην ανάλυση δικτύων

Μ. Αναγνώστου

23 Μαΐου 2022

Μοντέλο κίνησης με τη μορφή ροής

- Η κίνηση πάνω σ' ένα δίκτυο παριστάνεται συνήθως με ροές πάνω στις ακμές ενός γράφου.
- Οι κόμβοι του γράφου είναι οι κόμβοι του δικτύου και οι ακμές του γράφου είναι τα κανάλια του δικτύου.
- Η κίνηση εξομοιώνεται με αυτόν τον τρόπο με ένα υγρό. Αυτή η προσέγγιση βασίζεται σε όγκους δεδομένων και χωρητικότητες καναλιών.
- Αμελεί την ομαδοποίηση των bits σε πακέτα, τη χρήση πρωτοκόλλων, τα φαινόμενα αναμονής κ.α.

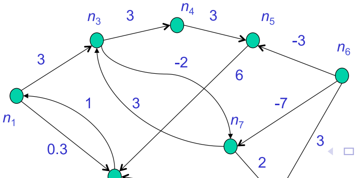


Διάνυσμα ροής

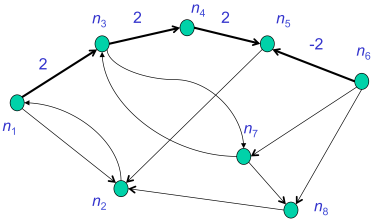
- Όλες μαζί οι ροές αποτελούν ένα *διάνυσμα ροής* x με συνιστώσες τις ροές x_{ij} επί των προσανατολισμένων ακμών (i, j).
- Στον γράφο του σχήματος το διάνυσμα ροής είναι

$$x = (x_{n_1 n_2}, x_{n_1 n_3}, x_{n_2 n_1}, x_{n_3 n_4}, x_{n_3 n_7}, \dots, x_{n_8 n_2}) = (0.3, 3, 1, 3, -2, \dots, -5)$$

- Η συνολικά εξερχόμενη ροή ενός κόμβου λέγεται *απόκλιση* (divergence) και προσδιορίζεται προσθέτοντας της ροές των εξερχομένων ακμών και αφαιρώντας τις ροές των εισερχομένων, π.χ. $d_{n_7} = x_{n_7 n_3} + x_{n_7 n_8} - x_{n_3 n_7} - x_{n_6 n_7} = 3 + 2 - (-2) - (-7) = 14$.
- Ένα κόμβος με θετική απόκλιση λέγεται *πηγή* (source), με αρνητική *καταβόθρα* (sink) και με μηδενική *κόμβος κυκλοφορίας* (circulation).



Απλή ροή μονοπατιού



- Η παχιά γραμμή δίνει ένα παράδειγμα απλής ροής μονοπατιού (simple path flow) ανάμεσα σε δύο κόμβους.
- Περιγράφει μια σταθερή ροή που εισάγεται στο δίκτυο από μια πηγή και εξαφανίζεται σε μια καταβόθρα.
- Μπορεί να είναι και κυκλική ροή (χωρίς πηγή και καταβόθρα).
- Η απλή ροή μονοπατιού *συμμορφώνεται* με μια δεδομένη συνολική ροή πάνω σε γράφο εφόσον «ρέει» προς την ίδια κατεύθυνση, δηλαδή είναι θετική (αρνητική) όπου είναι θετική (αρνητική) και η συνολική ροή.

Πραγματοποιήσιμη ροή

- Μια απλή ροή μονοπατιού θεωρείται πραγματοποιήσιμη και αντιστοιχεί σε μια μονόφορη επικοινωνία μεταξύ δύο άκρων. Εννοείται ότι το δίκτυο γνωρίζει τη σωστή δρομολόγηση.
- Ομοίως πραγματοποιήσιμη θεωρείται η υπέρθεση τέτοιων ρευμάτων.
- Πρόβλημα: Αν δοθεί ένα αυθαίρετο διάνυσμα ροής x (δηλαδή αυθαίρετες ροές ανά ακμή), είναι αυτή η ροή πραγματοποιήσιμη; (Δηλαδή μπορεί να αναλυθεί σε απλές ροές;)

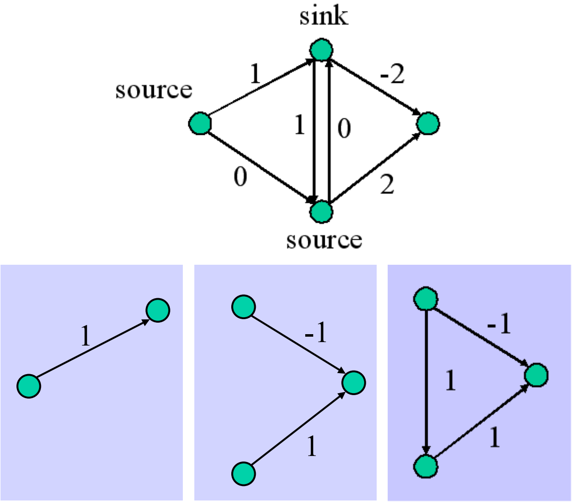
Θεώρημα της σύμμορφης αποσύνθεσης

- Ένα μη μηδενικό διάνυσμα ροής x μπορεί να αποσυντεθεί σε άθροισμα t απλών ροών μονοπατιού $x^1, x^2, \ldots, x^t$, δηλαδή

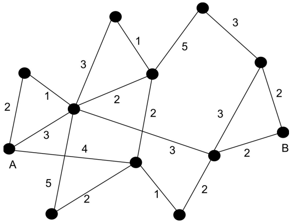
$$x = x^1 + x^2 + \ldots + x^t$$

  όπου κάθε απλή ροή x^i είναι ένα διάνυσμα ροής με ίσες τιμές για τις ακμές όπου αυτή εμφανίζεται και μηδενικές τιμές αλλού και τέτοιες ώστε να να συμμορφώνονται με την x (βλ. [0]).
- Οι απλές ροές ξεκινούν από πηγές της συνολικής ροής και καταλήγουν σε καταβόθρες της συνολικής ροής, εκτός αν στην συνολική ροή δεν υπάρχουν πηγές και καταβόθρες, οπότε οι απλές ροές είναι κυκλικές («κυκλοφορίες»).
- Αυτό σημαίνει ότι όλες οι ροές εν τέλει μπορούν να αποσυντεθούν σε απλές ροές, δηλαδή είναι «πραγματοποιήσιμες».

Παράδειγμα



Δρομολόγηση ελαχίστου κόστους



- Δίνεται ένα δίκτυο, όπου οι επιγραφές υποδηλώνουν κόστος μεταφοράς ανά μονάδα κίνησης.
- Αν επιλεγούν δύο κόμβοι ως πηγή-προορισμός, π.χ. A, B, κι ένα συνολικό μέγεθος φορτίου (π.χ. 100 μονάδες), πώς θα επιτευχθεί το ελάχιστο κόστος μεταφοράς; Πόσο είναι;
- Στη συνέχεια έστω ότι κάθε ακμή δεν μπορεί να μεταφέρει πάνω από 40 μονάδες κίνησης. Τι θα αλλάξει;

Δρομολόγηση ελαχίστου κόστους για ένα είδος κίνησης

- Έστω ότι έχουμε συγχρόνως πολλές πηγές και προορισμούς για το ίδιο είδος κίνησης (ένα «αγαθό», single commodity flow problem).
- Για παράδειγμα, έχουμε servers για ένα video streaming service, μελετάμε μόνο το download.
- Έχουμε προσδιορίσει την κατανομή φορτίου κάθε server (που αποτελεί μια πηγή) και το φορτίο που θα απορροφήσει κάθε client (καταβόθρα), ο οποίος μπορεί να συνδέεται με ένα ή περισσότερους servers.
- Πώς βρίσκουμε την πιο οικονομική ροή;
- Τι πρέπει να ισχύει για τις αποκλίσεις των πηγών και προορισμών συνολικά;
- Πώς θα βρούμε ποιος client συνδέεται με ποιον server (ή ποιους servers);

Το πρόβλημα της βέλτιστης ροής για ένα μεταφερόμενο αγαθό (single commodity flow problem)

Να υπολογισθεί το διάνυσμα ροής $\underline{x} = (x_{ij})$ που ελαχιστοποιεί τη συνάρτηση

$$\sum_{(i,j)\in\mathcal{A}} a_{ij}x_{ij}$$

με τους περιορισμούς

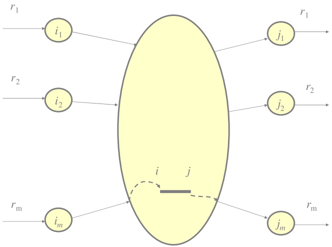
$$\sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij} - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} = s_i, \forall i \in \mathcal{N} \qquad (1)$$

και

$$b_{ij} \le x_{ij} \le c_{ij}, \quad \forall (i,j) \in \mathcal{A} \qquad (2)$$

όπου $\mathcal{A}$ είναι το σύνολο των ακμών,
a_{ij} το κόστος της ακμής (i,j),
b_{ij}, c_{ij} το κάτω και το άνω όριο ροής στην (i,j),
s_i η παροχή του κόμβου i και ικανοποιείται η συνθήκη $\sum_i s_i = 0$.
Ο περιορισμός (1) λέγεται περιορισμός *διατήρησης της ροής*.
Ο περιορισμός (2) λέγεται περιορισμός *χωρητικότητας*.

Προβλήματα ταυτόχρονης ροής διαφορετικών αγαθών (multicommodity flow problems)



Διαφορετικοί τύποι ροής μεταφέρονται μέσα από το ίδιο δίκτυο και εμφανίζουν αμοιβαία αλληλεπίδραση

- είτε επειδή μοιράζονται τους ίδιους αγωγούς
- είτε επειδή συμπλέκονται μέσω της αντικειμενικής συνάρτησης.

Παράδειγμα: Δρομολόγηση σε δίκτυα υπολογιστών I

Σκοπός είναι στο σχήμα της προηγούμενης διαφάνειας να γίνει δρομολόγηση κάθε ροής ξεχωριστά που φεύγει από τον κόμβο-πηγή i_k στον κόμβο-προορισμό j_k, $(k = 1, 2, \ldots, m)$ έτσι ώστε να ελαχιστοποιείται το συνολικό κόστος.

- Αν $x_{ij}(k)$ είναι η ροή τύπου k στο κανάλι (i, j), πρέπει να ισχύει για κάθε $k = 1, 2, \ldots, m$

$$\sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}(k) - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji}(k) =$$

$$= \begin{cases} r_k & \text{αν } i = i_k \\ -r_k & \text{αν } i = j_k \\ 0 & \text{αλλού} \end{cases}$$

- Οι ροές $x_{ij}(k)$ συνήθως πρέπει επίσης να είναι μη αρνητικές και να ικανοποιούν περιορισμούς μέγιστης ροής.

Παράδειγμα: Δρομολόγηση σε δίκτυα υπολογιστών II

- Το συνολικό κόστος προς ελαχιστοποίηση είναι συχνά της μορφής

$$f(x) = \sum_{(i,j)\in\mathcal{A}} f_{ij}(y_{ij})$$

  όπου f_{ij} είναι συνάρτηση της συνολικής ροής y_{ij}, που περνάει μέσα από τον αγωγό (i, j)

$$y_{ij} = \sum_{k=1}^{m} x_{ij}(k)$$

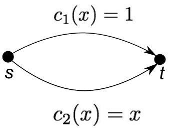
- Για παράδειγμα, συχνά θεωρούμε ότι το f παριστάνει μέση καθυστέρηση, η οποία όμως μπορεί να δίνεται από μη γραμμικό τύπο ως προς την κίνηση και να προκύπτει τελικά μη γραμμική συνάρτηση κόστους.

Αλγόριθμοι επίλυσης του προβλήματος της ροής

Θα αναφερθούν περαιτέρω τρεις κατηγορίες αλγορίθμων, που βασίζονται αντίστοιχα στις ακόλουθες ιδέες:

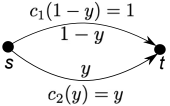
- *Βελτίωση του κόστους:* Κατασκευάζεται μια ακολουθία εφικτών ροών που βελτιώνει βήμα προς βήμα το κόστος.
- *Βελτίωση του δυϊκού κόστους:* Κατασκευάζεται το δυϊκό πρόβλημα με μεταβλητές λεγόμενες *τιμές*. Κατόπιν βελτιώνεται και πάλι βήμα προς βήμα το κόστος αλλάζοντας τις ροές.
- *Πλειστηριασμός:* Κατασκευάζεται μια ακολουθία τιμών, που υπενθυμίζει τη διαδικασία πλειστηριασμού.

Εγωιστική δρομολόγηση (Pigou, 1920)



- Δύο δρόμοι ενώνουν τις πόλεις *s*, *t*.
- Η καθυστέρηση (ή κόστος κ.λπ.) $c_1(x)$ στον πάνω δρόμο είναι σταθερή: $c_1(x) = 1$
- Η καθυστέρηση $c_2(x)$ στον κάτω δρόμο είναι ανάλογη της κίνησης: $c_2(x) = x$
- Η συνολικά διαθέσιμη κίνηση είναι ίση με 1, οπότε ποια κατανομή της ελαχιστοποιεί τη μέση καθυστέρηση;
- Τι θα συμβεί αν κάθε οδηγός αποφασίσει μόνος του «εγωιστικά»;

Συγκεντρωτική βέλτιστη λύση

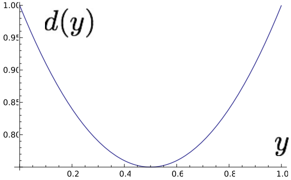


- Αν στον πάνω δρόμο η κίνηση είναι $1-y$, στον κάτω θα είναι y, οπότε η μέση καθυστέρηση είναι

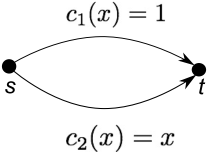
$$d(y) = (1-y) \times 1 + y \times y = 1 - y + y^2$$

- Επειδή $d'(y) = -1 + 2y$, η παράγωγος μηδενίζεται για $y = 1/2$, άρα το βέλτιστο είναι

$$d(1/2) = 1 - 1/2 + (1/2)^2 = 3/4$$

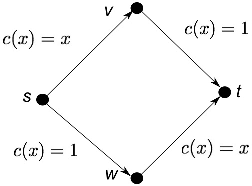


Εγωιστική λύση - Price of Anarchy



- Όλοι θα επιλέξουν τον κάτω δρόμο, σκεπτόμενοι ότι ακόμη κι αν ελάχιστοι άλλοι επιλέξουν τον πάνω δρόμο θα έχουν κέρδος.
- Άρα όλοι έχουν καθυστέρηση $c_2(1) = 1$.
- Το αποτέλεσμα είναι ότι κανείς δεν έχει κέρδος.
- Price of Anarchy = 4/3 (βλ. και [0]).

Το παράδοξο του Braess (1968): Αρχικό δίκτυο

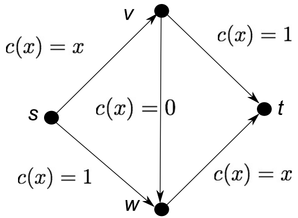


- Δύο δρόμοι ενώνουν τις πόλεις s, t.
- Η καθυστέρηση και στους δύο δρόμους είναι ίδια, ίση με $1 + x$.
- Αν στον πάνω δρόμο η κίνηση είναι z, στον κάτω θα είναι $1 - z$, οπότε η μέση καθυστέρηση είναι

$$d(y) = z \times (z + 1) + (1 - z) \times (1 - z + 1)$$

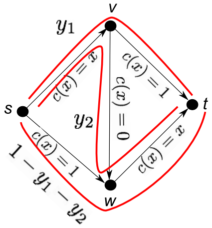
  και ελαχιστοποιείται σε 3/2 για $z = 1/2$.
- Το ίδιο αποτέλεσμα δίνει η εγωιστική δρομολόγηση.

Το παράδοξο του Braess: Βελτίωση του δικτύου



- Κατασκευάζουμε ένα δρόμο μηδενικής καθυστέρησης μεταξύ v, w.
- Το ερώτημα είναι τι αλλάζει στη βέλτιστη λύση και τι στην εγωιστική λύση.

Το παράδοξο του Braess: Βέλτιστη λύση



- Το κόστος ανά ακμή είναι:
  - Στο (s, v): $(y_1 + y_2)^2$
  - Στο (v, t): y_1
  - Στο (s, w): $1 - y_1 - y_2$
  - Στο (w, t): $(1 - y_1 - y_2) + y_2 = 1 - y_1$
  - Στο (v, w): 0
- Άρα το συνολικό κόστος είναι

$$f(y_1, y_2) = (y_1 + y_2)^2 + 1 - y_2 + (1 - y_1)^2$$

Βέλτιστη λύση

- Η συνάρτηση

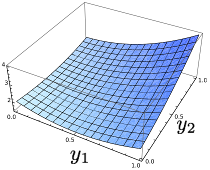
$$f(y_1, y_2) = (y_1 + y_2)^2 + 1 - y_2 + (1 - y_1)^2$$

με τους περιορισμούς $y_1 \geq 0, y_2 \geq 0$ ελαχιστοποιείται για

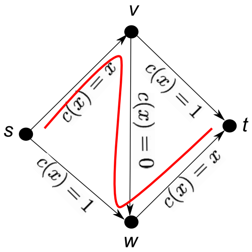
$$y_1 = 1/2, \quad y_2 = 0$$

και

$$\min_{y_1, y_2} f(y_1, y_2) = f(1/2, 0) = 3/2$$



Το παράδοξο του Braess: Εγωιστική λύση



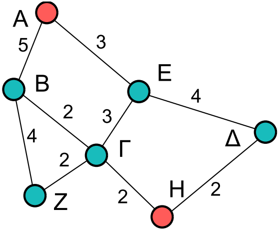
- Στην εγωιστική δρομολόγηση όλοι θα προτιμήσουν προφανώς τη διαδρομή $s - v - w - t$.
- Επομένως όλοι ($x = 1$) θα υποφέρουν καθυστέρηση ίση με

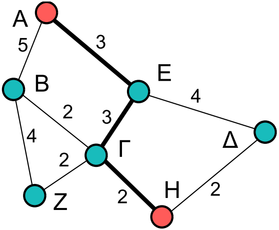
$$2x = 2$$

- Άρα η βελτίωση στο δίκτυο δίνει υποβάθμιση στην περίπτωση της εγωιστικής δρομολόγησης.

Ελάχιστο μονοπάτι

Αν οι επιγραφές σημαίνουν αποστάσεις, ποιο είναι το ελάχιστο μονοπάτι Α-Η;



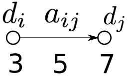


- Επιλύεται σε πολυωνυμικό χρόνο.
- Οι «αποστάσεις» μπορούν να παριστάνουν καθυστερήσεις, κόστη κ.λπ.

Αλγόριθμοι ελάχιστου μονοπατιού

Μια ευρεία κατηγορία αλγορίθμων χρησιμοποιεί ένα διάνυσμα «αποστάσεων» $(d_1, d_2, \ldots, d_N)$, του οποίου η συνιστώσα d_i αντιστοιχεί στον κόμβο i. Στηρίζονται στο εξής λήμμα:

- Έστω ότι τα (μονόμετρα) μεγέθη $d_1, d_2, \ldots, d_N$ ικανοποιούν $\forall (i,j) \in \mathcal{A}$ την ανισότητα $d_j \leq d_i + a_{ij}$, π.χ.



- Ας υποτεθεί ότι P είναι ένα μονοπάτι που αρχίζει από τον κόμβο i_1 και τελειώνει στον κόμβο i_k.

- Αν για όλες τις ακμές (i,j) του P ισχύει $d_j = d_i + a_{ij}$, το μονοπάτι P είναι ελάχιστο μεταξύ i_1 και i_k.

Απόδειξη του λήμματος

$$i_1 \longrightarrow i_2 \longrightarrow \cdots \longrightarrow i_{k-1} \longrightarrow i_k$$

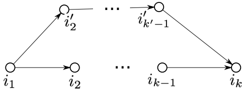
Έστω ότι η διαδρομή P περιλαμβάνει τους κόμβους $i_1, i_2, \ldots, i_k$. Προσθέτοντας κατά μέλη τις

$$\begin{aligned} d_{i_2} &= d_{i_1} + a_{i_1,i_2} \\ d_{i_3} &= di_2 + a_{i_2,i_3} \\ &\cdots \\ d_{i_k} &= d_{i_{k-1}} + a_{i_{k-1},i_k} \end{aligned}$$

προκύπτει ότι

$$d_{i_k} - d_{i_1} = \sum_{n=1}^{k-1} a_{i_n,i_{n+1}}.$$

Άρα για το μήκος $d(P) \equiv \sum_{n=1}^{k-1} a_{i_n,i_{n+1}}$ της διαδρομής P ισχύει $d(P) = d_{i_k} - d_{i_1}$.



Έστω άλλη διαδρομή $P' = (i_1, i_2', \ldots, i_{k'-1}', i_k)$.
Προσθέτοντας κατά μέλη τις

$$\begin{array}{rcl} d_{i_2'} & \leq & d_{i_1} + a_{i_1, i_2'} \\ d_{i_3'} & \leq & d_{i_2'} + a_{i_2', i_3'} \\ & \ldots & \\ d_{i_k} & \leq & d_{i_{k'-1}'} + a_{i_{k'-1}', i_k} \end{array}$$

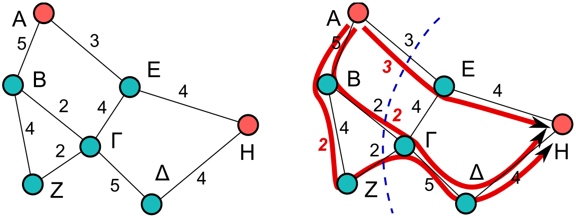
προκύπτει ότι

$$d_{i_k} - d_{i_1} \leq a_{i_1, i_2'} + \sum_{n=2}^{k'-2} a_{i_n', i_{n+1}'} + a_{i_{k'-1}', i_k}$$

Άρα για το μήκος της διαδρομής P' ισχύει $d(P) = d_{i_k} - d_{i_1} \leq d(P')$.

Μεταφορική ικανότητα: Μέγιστη ροή - ελάχιστη τομή

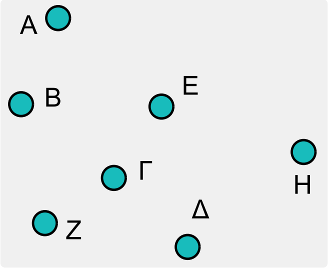
Στο σχήμα οι επιγραφές υποδηλώνουν χωρητικότητες. Πόση κίνηση μπορούμε να μεταφέρουμε απ' το Α στο Η;

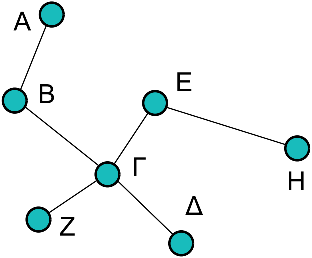


- Η μέγιστη ροή περιορίζεται απ' την *ελάχιστη τομή*.
- Η μέγιστη ροή υπολογίζεται με αλγόριθμο βαθμιαίας αύξησης της ροής (πολυωνυμικής πολυπλοκότητας).

Δίκτυο ελάχιστου μήκους: Ελάχ. διατρέχον δέντρο

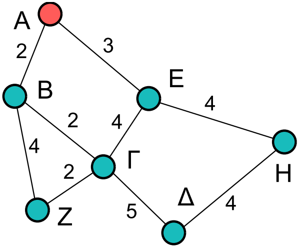
- Δίνονται κόμβοι με γνωστές αποστάσεις. Ποιο είναι το ελάχιστου συνολικού μήκους δίκτυο;
- Αλγόριθμος του Prim (greedy)

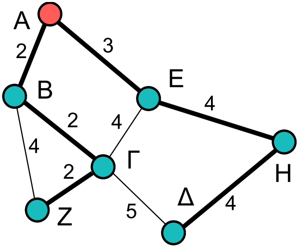




Εκπομπή: Ελάχιστου κόστους δέντρο διανομής

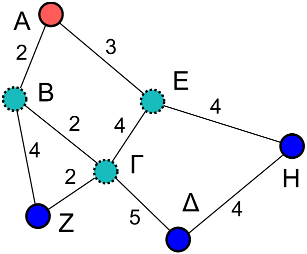
- Δίνεται δίκτυο με κόστος χρήσης ανά ζεύξη. Ποιο είναι το ελάχιστου συνολικού κόστους δίκτυο εκπομπής προερχόμενης από τον κόμβο Α;

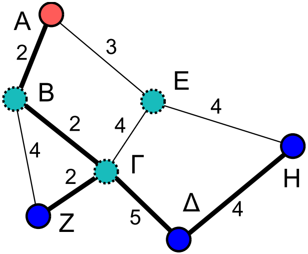




Πολυδιανομή: Ελάχιστου κόστους δέντρο διανομής

Δίνεται δίκτυο με κόστος χρήσης ανά ζεύξη. Ποιο είναι το ελάχιστου συνολικού κόστους δίκτυο διανομής από τον κόμβο Α προς τους κόμβους Ζ, Δ, Η;





- Το βέλτιστο δέντρο είναι γνωστό ως δέντρο Steiner.
- Πρόκειται για δύσκολο (NP-πλήρες) πρόβλημα.

Βιβλιογραφία

D. P. Bertsekas. *Network optimization: continuous and discrete models*. Athena Scientific, 1998.

T. Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005. ISBN: 0262182432.

Τελευταία ενημέρωση: 23 Μαΐου 2022 07:57