Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Αλγόριθμοι και Πολυπλοκότητα
4η Σειρά Προτεινόμενων Ασκήσεων

Άσκηση 1: Βελτίωση Οδικού Δικτύου (DPV 4.10)

Έστω οδικό δίκτυο $G(V, E, \ell)$ που συνδέει ένα σύνολο πόλεων V. Θεωρούμε ότι το δίκτυο είναι κατευθυνόμενο, και ότι κάθε δρόμος $(u, v) \in E$ έχει (μη αρνητικό) μήκος $\ell(u, v)$. Πρόκειται να κατασκευαστεί ένας νέος δρόμος, και υπάρχει λίστα E' με ακμές-προτάσεις για ζεύγη πόλεων τα οποία μπορεί να συνδέσει. Κάθε ακμή-πρόταση $(u, v) \in E'$ συνοδεύεται από αντίστοιχο μήκος $\ell'(u, v)$. Το ζητούμενο είναι να επιλέξουμε την ακμή-πρόταση που επιτυγχάνει την μέγιστη μείωση της απόστασης μεταξύ δύο δεδομένων πόλεων $s, t \in V$. Να διατυπώσετε έναν όσο το δυνατόν πιο αποδοτικό αλγόριθμο για αυτό το πρόβλημα. Να αιτιολογήσετε την ορθότητα και την πολυπλοκότητα του αλγορίθμου σας.

Λύση. Μια πρώτη ιδέα που θα μπορούσαμε να δοκιμάσουμε είναι για κάθε νέο δρόμο (u, v), να κατασκευάσουμε ένα νέο δίκτυο G', όπου προσθέτουμε την ακμή (u, v) στο αρχικό δίκτυο G. Για κάθε τέτοιο δίκτυο G', μπορούμε να βρούμε το συντομότερο $s - t$ μονοπάτι με τον αλγόριθμο του Dijkstra. Τέλος, αρκεί να συγκρίνουμε το συντομότερο $s - t$ μονοπάτι από όλα τα δίκτυα G' με το συντομότερο $s - t$ μονοπάτι στο δίκτυο G. Αυτή η λύση χρειάζεται χρόνο $O(|E'|(|E| + |V| \log |V|))$.

Ο παραπάνω αλγόριθμος δεν είναι ο πιο αποδοτικός. Η ιδέα που χρειαζόμαστε είναι ότι ένα μονοπάτι το οποίο μειώνει το κόστος πρέπει να χρησιμοποιεί ακριβώς μία από τις νέες ακμές του συνόλου E' (καθώς μπορούμε να "αγοράσουμε" μόνο μια ακμή). Συνεπώς, αρκεί για κάθε ακμή (u, v), να εξετάσουμε το συντομότερο μονοπάτι που ξεκινά από τον κόμβο s, περνάει από την ακμή (u, v) και καταλήγει στον κόμβο t. Για να υπολογιστεί αυτό το μονοπάτι, χρειαζόμαστε μόνο το συντομότερο μονοπάτι από το s στο u και το συντομότερο μονοπάτι από το v στο t στο δίκτυο G. Το κόστος θα είναι τότε $\mathrm{dist}(s, u) + \ell'(u, v) + \mathrm{dist}(v, t)$.

Μπορούμε να υπολογίσουμε τις αποστάσεις που χρειαζόμαστε με δύο εκτελέσεις του αλγορίθμου του Dijkstra. Αρχικά εκτελούμε τον Dijkstra στο δίκτυο G με αρχική κορυφή s, και υπολογίζουμε τις αποστάσεις $\mathrm{dist}(s, w)$, για κάθε $w \in V$. Για να υπολογίσουμε τις αποστάσεις $\mathrm{dist}(w, t)$, $w \in V$, εκτελούμε τον Dijkstra με αρχική κορυφή t στο δίκτυο που προκύπτει από το G αντιστρέφοντας τη φορά κάθε ακμής. Τέλος, υπολογίζουμε για κάθε ακμή $(u, v) \in E'$, το άθροισμα $\mathrm{dist}(s, u) + \ell'(u, v) + \mathrm{dist}(v, t)$, βρίσκουμε το ελάχιστο τέτοιο άθροισμα, και το συγκρίνουμε με την απόσταση $s - t$ στο αρχικό δίκτυο G. Η χρονική πολυπλοκότητα του αλγορίθμου είναι $\mathrm{O}(|E'| + |E| + |V| \log |V|)$. $\square$

Άσκηση 2: Αραιά Γραφήματα με Σύντομα Μονοπάτια (KT 4.31)

Έστω μη κατευθυνόμενο γράφημα $G(V, E, w)$ με θετικά μήκη στις ακμές, και έστω $d_G(u, v)$ η απόσταση των κορυφών u και v στο G με βάση τα μήκη w. Εφαρμόζουμε στο G την παρακάτω παραλλαγή του αλγόριθμου του Kruskal με στόχο να κατασκευάσουμε ένα αραιό συνδετικό υπογράφημα $H(V, E')$, $E' \subseteq E$, όπου η απόσταση $d_H(u, v)$ προσεγγίζει ικανοποιητικά την απόσταση $d_G(u, v)$.

Εξετάζουμε διαδοχικά τις ακμές του G σε αύξουσα σειρά του μήκους τους (αν σας βοηθάει, μπορείτε να υποθέσετε ότι όλα τα μήκη των ακμών είναι διαφορετικά). Η τρέχουσα ακμή $e = \{u, v\}$ προστίθεται στο H αν είτε το H δεν περιέχει $u-v$ μονοπάτι είτε το συντομότερο $u - v$ μονοπάτι που περιέχει το H έχει μήκος *μεγαλύτερο* από $3w(e)$.

(α) Να δείξετε ότι για κάθε ζευγάρι κορυφών u, v, ισχύει ότι $d_H(u, v) \leq 3d_G(u, v)$, όπου $d_H(u, v)$ η απόσταση των u και v στο γράφημα H που υπολογίζεται από τον παραπάνω αλγόριθμο.

(β) Να δείξετε ότι κάθε απλό μη κατευθυνόμενο γράφημα με n κορυφές και $\omega(n^{3/2})$ ακμές περιέχει κύκλο με μήκος μικρότερο ή ίσο του 4. *Σημείωση:* Γενικότερα, ισχύει ότι κάθε απλό μη κατευθυνόμενο γράφημα με n κορυφές και $\omega(n^{1+1/k})$ ακμές περιέχει κύκλο με μήκος μικρότερο ή ίσο του $2k$.

(γ) Να δείξετε ότι το πλήθος ακμών του H είναι $O(n^{3/2})$, όπου n το πλήθος κορυφών του G.

Λύση. (α) Έστω $e = \{x, y\}$ μια οποιαδήποτε ακμή του G. Αρχικά παρατηρούμε ότι $d_H(x, y) \leq 3w(e)$. Πράγματι, αν η e ανήκει στον H, τότε $d_H(x, y) \leq w(e) \leq 3w(e)$. Αν η e δεν ανήκει στον H, τότε από την περιγραφή του αλγόριθμου, $d_H(x, y) \leq 3w(e)$.

Αν λοιπόν θεωρήσουμε ένα οποιοδήποτε ζεύγος κορυφών u, v και ένα συντομότερο $u - v$ μονοπάτι $p = (u = x_0, x_1, \ldots, x_{k-1}, x_k = v)$, έχουμε ότι

$$3d_G(u, v) = \sum_{i=0}^{k-1} 3w(\{x_i, x_{i+1}\}) \geq \sum_{i=0}^{k-1} d_H(x_i, x_{i+1}) \geq d_H(u, v)$$

(β) Έστω ένα οποιοδήποτε (απλό μη κατευθυνόμενο) γράφημα G με n κορυφές που δεν έχει κύκλο μήκους 3 ή 4. Ένα άνω φράγμα στον αριθμό των ακμών του G προκύπτει υπολογίζοντας το πλήθος των διαφορετικών "κερασιών" $K_{1,2}$ (δηλ. επαγόμενων υπογραφημάτων με 3 κορυφές εκ των οποίων οι δύο "ακραίες" συνδέονται με ακμή με την "κεντρική" αλλά όχι μεταξύ τους) με δύο διαφορετικούς τρόπους, και συγκρίνοντας τα αποτελέσματα.

Κάθε κορυφή v βαθμού $d(v)$ αποτελεί το "κέντρο" $\binom{d(v)}{2}$ διαφορετικών "κερασιών". Συνεπώς το πλήθος των διαφορετικών "κερασιών" στον G είναι ίσο με:

$$\sum_{v \in V} \binom{d(v)}{2}$$

Επίσης παρατηρούμε ότι το πλήθος των "κερασιών" του G δεν ξεπερνά το $\binom{n}{2}$ (δηλ. το πλήθος των ζευγαριών κορυφών που μπορούν να αποτελέσουν τα "άκρα" ενός "κερασιού"). Έτσι καταλήγουμε ότι:

$$\binom{n}{2} \geq \sum_{v \in V} \binom{d(v)}{2} \geq n \binom{\overline{d}}{2},$$

όπου $\overline{d} = 2m/n$ είναι ο μέσος βαθμός του G, και η τελευταία ανισότητα ισχύει γιατί η συνάρτηση $\binom{x}{2}$ είναι κυρτή ως προς x. Λύνοντας ως προς $\overline{d}$, προκύπτει ότι $\overline{d} \leq \sqrt{n-1} + 1$, δηλ. ότι

$$m \leq \frac{n\sqrt{n-1} + n}{2}$$

(γ) Θα δείξουμε ότι ο H δεν περιέχει κύκλους μήκους 3 ή 4, από το οποίο, λόγω του (β), προκύπτει το ζητούμενο. Στη συνέχεια αποκλείουμε την ύπαρξη κύκλου μήκους 4 στον H. Η ύπαρξη κύκλου μήκους 3 στον H αποκλείεται με τον ίδιο τρόπο.

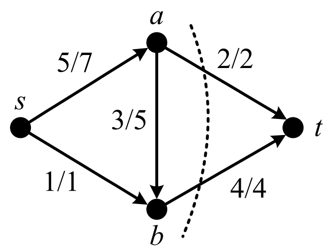
Έστω ότι υπάρχει στον H κύκλος (x, y, z, w) μήκους 4, και έστω $e_1 = \{x, y\}, e_2 = \{y, z\}, e_3 = \{z, w\}$, και $e_4 = \{w, x\}$ οι αντίστοιχες ακμές. Χωρίς βλάβη της γενικότητας, θεωρούμε ότι $w(e_1) \leq w(e_2) \leq w(e_3) \leq w(e_4)$, και επομένως η e_4 εξετάζεται για προσθήκη στον H, αφού έχουν ήδη επιλεγεί οι ακμές e_1, e_2, και e_3. Την στιγμή που ο αλγόριθμος εξετάζει την e_4, ισχύει για την απόσταση $d_H(x, w)$ των άκρων της στον H ότι:

$$\begin{aligned} d_H(x,w) &\leq d_H(x,y) + d_H(y,z) + d_H(z,w) \\ &\leq w(e_1) + w(e_2) + w(e_3) \leq 3w(e_4) \end{aligned}$$

Επομένως η ακμή e_4 δεν επιλέγεται από τον αλγόριθμο, και ο αντίστοιχος κύκλος δεν μπορεί να σχηματιστεί στον H. □

Άσκηση 3: Μείωση Μεταφορικής Ικανότητας (ΚΤ 7.12)

Θεωρούμε δίκτυο ροής $G(V, E)$ με αρχική κορυφή $s \in V$, καταληκτική κορυφή $t \in V$, και μοναδιαίες χωρητικότητες στις ακμές, και φυσικό αριθμό k. Το ζητούμενο είναι να βρούμε k ακμές η διαγραφή των οποίων προκαλεί την μεγαλύτερη δυνατή μείωση στην μέγιστη ροή μεταξύ των κορυφών s και t. Να διατυπώσετε έναν αποδοτικό αλγόριθμο για αυτό το πρόβλημα. Να αιτιολογήσετε την ορθότητα του αλγορίθμου σας, και να διερευνήσετε αν η προσέγγιση που ακολουθήσατε γενικεύεται για ακμές διαφορετικής χωρητικότητας.



Σχήμα 1. Παράδειγμα δικτύου με διαφορετικές χωρητικότητες. Σε κάθε ακμή σημειώνεται η ροή που την διατρέχει στην μέγιστη $s-t$ ροή και η χωρητικότητά της. Η ελάχιστη $s-t$ τομή είναι η $(\{s, a, b\}, \{t\})$ και σημειώνεται με την διακεκομμένη γραμμή. Παρατηρούμε ότι η διαγραφή της ακμής (b, t) μειώνει την μέγιστη ροή κατά 4, ενώ η διαγραφή της ακμής (s, a) μειώνει την μέγιστη ροή κατά 5.

Λύση. Η μεγαλύτερη δυνατή μείωση στην μέγιστη ροή προκύπτει αν διαγράψουμε οποιεσδήποτε k ακμές από την ελάχιστη $s - t$ τομή. Πράγματι, αφού όλες οι ακμές είναι μοναδιαίας χωρητικότητας, η διαγραφή μιας ακμής μειώνει την μέγιστη ροή το πολύ κατά 1. Η διαγραφή k ακμών από την ελάχιστη τομή μειώνει την ελάχιστη τομή (και συνεπώς και την μέγιστη ροή) κατά k. Η ελάχιστη τομή υπολογίζεται π.χ. με τον αλγόριθμο Edmonds-Karp σε χρόνο $\mathrm{O}(|V||E|^2)$.

Και στην γενική περίπτωση, όπου οι ακμές έχουν διαφορετικές χωρητικότητες, διαγράφοντας μια ακμή της ελάχιστης τομής έχουμε μείωση της μέγιστης ροής ίση με την χωρητικότητα της ακμής. Όπως όμως φαίνεται στο Σχήμα 1, αυτό δεν είναι απαραίτητα το καλύτερο δυνατό. □

Άσκηση 4: Προγραμματισμός Εφημεριών (ΚΤ 7.19)

Ένα νοσοκομείο που απασχολεί k γιατρούς προγραμματίζει τις εφημερίες τους για τις επόμενες n ημέρες. Κάθε γιατρός j, $1 \leq j \leq k$, έχει δηλώσει ένα σύνολο ημερών L_j στις οποίες είναι διαθέσιμος για εφημερία, και ότι ο μέγιστος αριθμός εφημεριών που μπορεί να κάνει είναι ℓ_j. Επιπλέον, για κάθε ημέρα i, $1 \leq i \leq n$, είναι γνωστός ο ακριβής αριθμός p_i των γιατρών που εφημερεύουν. Πρέπει λοιπόν να υπολογισθεί ένα σύνολο ημερών εφημερίας $L'_j \subseteq L_j$, με $|L'_j| \leq \ell_j$, για κάθε γιατρό j, ώστε κάθε ημέρα i να περιλαμβάνεται στα σύνολα ημερών εφημερίας ακριβώς p_i γιατρών, ή να διαπιστωθεί ότι κάτι τέτοιο δεν είναι εφικτό. Να διατυπώσετε έναν αποδοτικό αλγόριθμο για αυτό το πρόβλημα. Να αιτιολογήσετε την ορθότητα και την πολυπλοκότητα του αλγορίθμου σας.

Λύση. Η λύση του προβλήματος ανάγεται στην εύρεση μιας μέγιστης ροής σε ένα κατάλληλα διαμορφωμένο δίκτυο $G(V, E)$ που αναπαριστά τους περιορισμούς του προβλήματος. Το δίκτυο G κατασκευάζεται ως εξής:

1. Θεωρούμε μια αρχική κορυφή s, μια καταληκτική κορυφή t, ένα σύνολο κορυφών V_1 που περιέχει μια κορυφή για κάθε γιατρό, και ένα σύνολο κορυφών V_2 που περιέχει μια κορυφή για κάθε ημέρα. Το σύνολο κορυφών του δικτύου είναι $V = \{s\} \cup V_1 \cup V_2 \cup \{t\}$, με $|V| = n + k + 2$.
2. Η κορυφή s συνδέεται με ακμή χωρητικότητας ℓ_j με κάθε κορυφή $j \in V_1$.
3. Κάθε κορυφή $i \in V_2$ συνδέεται με ακμή χωρητικότητας p_i με την κορυφή t.
4. Κάθε κορυφή $j \in V_1$ συνδέεται με κάθε κορυφή $i \in L_j$ με ακμή μοναδιαίας χωρητικότητας.
5. Καμία άλλη ακμή δεν υπάρχει στο δίκτυο. Το πλήθος των ακμών του δικτύου είναι $\mathrm{O}(nk)$.

Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι $\sum_{j \in [k]} \ell_j \geq \sum_{i \in [n]} p_i$ (διαφορετικά δεν υπάρχει εφικτό πρόγραμμα εφημεριών τετριμμένα). Προφανώς η μέγιστη $s - t$ ροή στο δίκτυο G είναι μικρότερη ή ίση του $\sum_{i \in [n]} p_i$. Θα δείξουμε ότι υπάρχει αποδεκτό πρόγραμμα εφημεριών ανν η μέγιστη $s - t$ ροή στο δίκτυο G είναι ίση με $\sum_{i \in [n]} p_i$. Συγκεκριμένα, θα δείξουμε ότι κάθε εφικτό πρόγραμμα εφημεριών μπορεί να μετατραπεί σε μια $s - t$ ροή μεγέθους $\sum_{i \in [n]} p_i$, και αντίστροφα.

Αρχικά θεωρούμε ένα εφικτό πρόγραμμα εφημεριών $(L'_1, \ldots, L'_k)$. Σε κάθε ακμή (s, j) δρομολογούμε ροή ίση με $|L'_j|$, σε κάθε ακμή (j, i) δρομολογούμε μια μονάδα ροής αν $i \in L'_j$, και μηδενική ροή διαφορετικά, και σε κάθε ακμή (i, t) δρομολογούμε ροή $|\{j : i \in L'_j\}|$. Εξ' ορισμού, η ροή διατηρείται σε κάθε ενδιάμεσο κόμβο. Επιπλέον, αφού το πρόγραμμα εφημεριών $(L'_1, \ldots, L'_k)$ είναι εφικτό, πρέπει $|L'_j| \leq \ell_j$ και $|\{j : i \in L'_j\}| = p_i$. Άρα έχουμε μια εφικτή (και άρα μέγιστη) $s - t$ ροή μεγέθους $\sum_{i \in [n]} p_i$.

Αντίστροφα, θεωρούμε μια ακέραια $s - t$ ροή μεγέθους $\sum_{i \in [n]} p_i$ (από τον αλγόριθμο Ford-Fulkerson, γνωρίζουμε ότι κάθε δίκτυο με ακέραιες χωρητικότητες έχει μια ακέραια μέγιστη ροή). Άρα η ροή σε κάθε ακμή (j, i), $j \in V_1$, $i \in V_2$, είναι είτε 1 είτε 0. Αναθέτουμε σε έναν γιατρό j να εφημερεύει την ημέρα i ανν η ροή στην ακμή (j, i) είναι 1. Το γεγονός ότι το συγκεκριμένο πρόγραμμα εφημεριών είναι εφικτό προκύπτει εύκολα από τις ιδιότητες και το μέγεθος της ροής.

Η χρονική πολυπλοκότητα είναι αυτή που χρειάζεται για να λύσουμε το πρόβλημα της μέγιστης ροής σε ένα δίκτυο με $\Theta(n + k)$ κορυφές, $\mathrm{O}(nk)$ ακμές, και χωρητικότητες $\mathrm{O}(n + k)$. □

Άσκηση 5: Αναγωγές και NP-Πληρότητα

Να δείξετε ότι τα παρακάτω προβλήματα είναι **NP**-Πλήρη:

Συνδετικό Δέντρο με Ελάχιστο Αριθμό Φύλλων (Min Leaf Spanning Tree)

Είσοδος: Μη κατευθυνόμενο γράφημα $G(V, E)$ και φυσικός αριθμός k, $2 \leq k < |V|$.

Ερώτηση: Έχει το G συνδετικό δέντρο με k ή λιγότερα φύλλα (δηλ. κορυφές βαθμού 1);

Κυρίαρχο Σύνολο (Dominating Set)

Είσοδος: Μη κατευθυνόμενο γράφημα $G(V, E)$ και φυσικός αριθμός k, $1 \leq k \leq |V|$.

Ερώτηση: Έχει το G κυρίαρχο σύνολο $D \subseteq V$ με $|D| \leq k$; Ένα σύνολο κορυφών $D \subseteq V$ αποτελεί κυρίαρχο σύνολο για ένα μη κατευθυνόμενο γράφημα $G(V, E)$ αν κάθε κορυφή $v \in V$ είτε ανήκει στο D είτε συνδέεται με κάποια κορυφή του D.

Δέντρο Steiner (Steiner Tree)

Είσοδος: Πλήρες μη κατευθυνόμενο γράφημα $G(V,E)$ με μήκη $\ell : E \mapsto \mathbb{N}$ στις ακμές, υποσύνολο κορυφών $C \subset V$, και φυσικός αριθμός B. Τα μήκη των ακμών ικανοποιούν την τριγωνική ανισότητα, δηλ. για κάθε $x, y, z \in V$, $\ell(x,y) \leq \ell(x,z) + \ell(z,y)$.

Ερώτηση: Υπάρχει συνεκτικό υπογράφημα του G που περιλαμβάνει όλες τις κορυφές του C (και ενδεχομένως κάποιες κορυφές του $V \setminus C$) με συνολικό μήκος που δεν ξεπερνά το B;

Σημείωση: Μπορούμε να θεωρήσουμε ότι το ζητούμενο υπογράφημα είναι δέντρο. Ένα τέτοιο δέντρο ονομάζεται *δέντρο Steiner* για το C. Το πρόβλημα είναι διαφορετικό από τον υπολογισμό ενός Ελάχιστου Συνδετικού Δέντρου, αφού το βέλτιστο δέντρο Steiner δεν περιλαμβάνει κατ' ανάγκη όλες τις κορυφές του V, αλλά μπορεί να εκμεταλλευθεί την ύπαρξη κάποιων κορυφών του $V \setminus C$ για να επιτύχει μικρότερο συνολικό μήκος σε σχέση με το συνολικό μήκος ενός Ελάχιστου Συνδετικού Δέντρου στο επαγόμενο υπογράφημα που ορίζεται από το C.

Χωροθέτηση Υπηρεσιών (Facility Location)

Είσοδος: Πλήρες μη κατευθυνόμενο γράφημα $G(V,E)$ με αποστάσεις $\ell : E \mapsto \mathbb{N}$ στις ακμές και κόστη $f : V \mapsto \mathbb{N}$ στις κορυφές, και φυσικός αριθμός B. Οι αποστάσεις ικανοποιούν την τριγωνική ανισότητα, δηλ. για κάθε $x, y, z \in V$, $\ell(x,y) \leq \ell(x,z) + \ell(z,y)$.

Ερώτηση: Υπάρχει σύνολο κορυφών $C \subseteq V$ για το οποίο το συνολικό κόστος των κορυφών του C και η συνολική απόσταση κάθε κορυφής του $V \setminus C$ από την πλησιέστερη κορυφή στο C δεν ξεπερνά το B; Πρέπει δηλαδή να διαπιστώσουμε αν υπάρχει $C \subseteq V$ με συνολικό κόστος

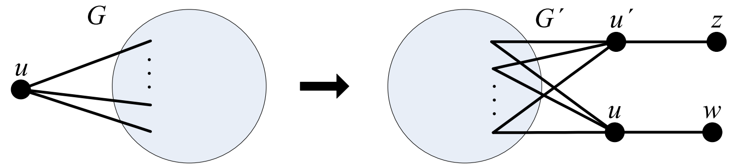
$$\sum_{v \in C} f(v) + \sum_{u \in V \setminus C} \min_{v \in C} \{\ell(u,v)\} \leq B \tag{1}$$

Σημείωση: Η αντιστοιχία με το πρόβλημα χωροθέτησης υπηρεσιών προκύπτει αν θεωρήσουμε ότι σε κάθε κορυφή $u \in V$ υπάρχει ένας πελάτης που πρέπει να συνδεθεί στην πλησιέστερη θέση όπου παρέχεται μια υπηρεσία. Οι αποστάσεις αντιστοιχούν στα κόστη σύνδεσης των πελατών. Το κόστος $f(v)$ αντιστοιχεί στο κόστος εγκατάστασης της υπηρεσίας στη θέση v. Επιλέγοντας να εγκαταστήσουμε την υπηρεσία στις κορυφές του C, έχουμε συνολικό κόστος εγκατάστασης $\sum_{v \in C} f(v)$, και κόστος σύνδεσης για κάθε κορυφή $u \in V \setminus C$ ίσο με $\min_{v \in C}\{\ell(u,v)\}$ (κάθε πελάτης συνδέεται στην πλησιέστερη θέση, το κόστος σύνδεσης για τους πελάτες στο C είναι μηδενικό). Το συνολικό κόστος του C δίνεται από την (1).

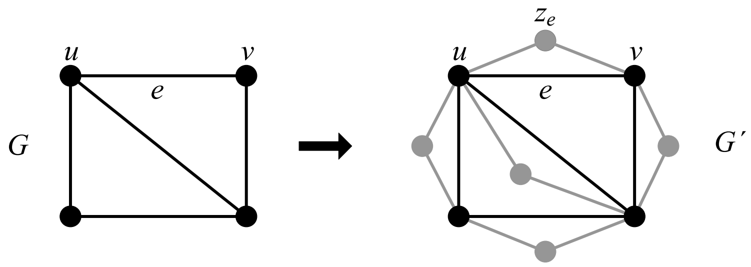
Υπόδειξη: Για το δεύτερο και το τρίτο πρόβλημα, μπορείτε να δοκιμάσετε αναγωγές από το Κάλυμμα Κορυφών (Vertex Cover). Για το τέταρτο πρόβλημα, μπορείτε να δοκιμάσετε αναγωγή από το Κυρίαρχο Σύνολο (Dominating Set).

Λύση. Όλα τα προβλήματα ανήκουν στο $\mathbf{NP}$, αφού οι υποψήφιες λύσεις τους μπορούν να ελεγχθούν σε πολυωνυμικό χρόνο. Στη συνέχεια παρουσιάζουμε αναγωγές από γνωστά $\mathbf{NP}$ πλήρη προβλήματα, αποδεικνύοντας ότι τα συγκεκριμένα προβλήματα είναι $\mathbf{NP}$-πλήρη.

Αναγωγή από το πρόβλημα του Κύκλου Hamilton στο πρόβλημα του Συνδετικού Δέντρου με Ελάχιστο Αριθμό Φύλλων. Έστω γράφημα $G(V,E)$ στο οποίο ελέγχουμε την ύπαρξη κύκλου Hamilton. Κατασκευάζουμε γράφημα G' από το G ως εξής: Θεωρούμε μια (αυθαίρετα επιλεγμένη) κορυφή $u \in V$, προσθέτουμε ένα "αντίγραφο" u' της u, και συνδέουμε την u' με όλους τους γείτονες της u στο G (και μόνο με αυτούς). Επιπλέον, προσθέτουμε κορυφή w, που συνδέεται μόνο με την u, και κορυφή z, που συνδέεται μόνο με την u' (βλ. Σχήμα 2). Θέτουμε ακόμη $k = 2$. Προφανώς το G' μπορεί να κατασκευαστεί από το G σε πολυωνυμικό χρόνο.



Σχήμα 2. Σχηματική αναπαράσταση της αναγωγής από το πρόβλημα του κύκλου Hamilton στο πρόβλημα του του Συνδετικού Δέντρου με Ελάχιστο Αριθμό Φύλλων.



Σχήμα 3. Ένα παράδειγμα για την αναγωγή από το πρόβλημα του Καλύμματος Κορυφών στο πρόβλημα του Κυρίαρχου Συνόλου. Οι ακμές και κορυφές που προστίθενται στο G' σημειώνονται γκρίζες.

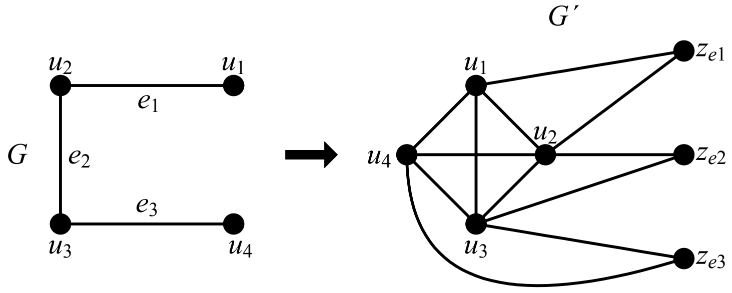
Θα δείξουμε ότι το G' έχει συνδετικό δέντρο με (το πολύ) 2 φύλλα ανν το G έχει κύκλο Hamilton. Έστω ότι το G έχει κύκλο Hamilton $u - p - u$, όπου p μονοπάτι που περιέχει όλες τις κορυφές του G εκτός της u. Τότε το G' περιέχει το μονοπάτι $w - u - p - u' - z$, το οποίο αποτελεί ένα συνδετικό δέντρο του G' με 2 φύλλα.

Αντίστροφα, έστω ότι το G' έχει συνδετικό δέντρο με (το πολύ) 2 φύλλα. Κατ΄ ανάγκη, αυτό θα είναι ένα μονοπάτι που περιέχει όλες τις κορυφές του G' και έχει άκρα τις κορυφές w και z. Συνεπώς το γράφημα G' περιέχει μονοπάτι $w - u - p - u' - z$, όπου p μονοπάτι που περιέχει όλες τις κορυφές του G' εκτός των u, u', w, z. Τότε ο κύκλος $u - p - u$ αποτελεί κύκλο Hamilton για το γράφημα G.

Παρατήρηση: Στο πρόβλημα του Μονοπατιού Hamilton, δίνεται ένα γράφημα G, και ελέγχουμε αν υπάρχει (απλό) μονοπάτι που διέρχεται από όλες τις κορυφές του G. Η παραπάνω κατασκευή ουσιαστικά αποτελεί μια αναγωγή από το πρόβλημα του Κύκλου Hamilton στο πρόβλημα του Μονοπατιού Hamilton. Έτσι αποδεικνύουμε ότι το πρόβλημα του Μονοπατιού Hamilton είναι **NP**-πλήρες. Το πρόβλημα του Συνδετικού Δέντρου με Ελάχιστο Αριθμό Φύλλων ουσιαστικά αποτελεί γενίκευση του Μονοπατιού Hamilton, και συνεπώς είναι και αυτό **NP**-πλήρες. □

Αναγωγή του Καλύμματος Κορυφών στο Κυρίαρχο Σύνολο. Έστω γράφημα $G = (V, E)$ στο οποίο ελέγχουμε αν υπάρχει κάλυμμα κορυφών C με $|C| \leq k$. Χωρίς βλάβη της γενικότητας, θεωρούμε ότι το γράφημα G είναι συνεκτικό (ως άσκηση, να αιτιολογήσετε αυτό το σημείο).

Κατασκευάζουμε γράφημα G' ώστε κάθε κυρίαρχο σύνολο του G' να αντιστοιχεί σε ένα κάλυμμα κορυφών του G, και αντίστροφα. Το G' προκύπτει από το G προσθέτοντας ένα "παράλληλο" μονοπάτι μήκους 2 για κάθε ακμή του G. Ειδικότερα, το G' περιέχει όλες τις κορυφές και τις ακμές του G. Επιπλέον, για κάθε ακμή $e = \{u, v\} \in E$, το G' περιέχει μια επιπλέον κορυφή z_e και δύο επιπλέον ακμές $\{u, z_e\}$ και $\{z_e, v\}$ που συνδέουν τις u και v με μονοπάτι μήκους 2 μέσω της z_e (βλ. Σχήμα 3). Προφανώς το G' μπορεί να κατασκευαστεί από το G σε πολυωνυμικό χρόνο.



Σχήμα 4. Ένα παράδειγμα για την αναγωγή από το πρόβλημα του Καλύμματος Κορυφών στο πρόβλημα του Δέντρου Steiner. Από τις ακμές του (πλήρους) γραφήματος G' εμφανίζονται μόνο αυτές με μήκος 1. Από τις ακμές του G' που δεν εμφανίζονται, η ακμή $\{e_1, e_3\}$ έχει μήκος 3, και όλες οι υπόλοιπες έχουν μήκος 2.

Θα δείξουμε ότι το γράφημα G έχει κάλυμμα κορυφών C με $|C| \leq k$ ανν το γράφημα G' έχει κυρίαρχο σύνολο D με $|D| \leq k$. Πράγματι, κάθε κάλυμμα κορυφών C του G αποτελεί κυρίαρχο σύνολο για το G' γιατί: (α) κάθε κορυφή $u \in V$, είτε ανήκει στο C είτε (εφόσον δεν είναι απομονωμένη) συνδέεται με κορυφή του C (στο G, και άρα στο G'), και (β) κάθε κορυφή z_e συνδέεται με κάποια κορυφή του C στο G' (αφού η αντίστοιχη ακμή e "καλύπτεται" από το C στο G).

Αντίστροφα, έστω D ένα κυρίαρχο σύνολο στο G'. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι $D \subseteq V$ (διαφορετικά, κάθε $z_e \in D$, όπου $e = \{u, v\}$, "κυριαρχεί" μόνο στις z_e, u, v, έτσι μπορούμε να αντικαταστήσουμε την z_e με μία από τις u, v στο D). Κάθε κορυφή z_e συνδέεται με κάποια κορυφή του D στο G'. Επομένως κάθε ακμή $e \in E$ του G έχει τουλάχιστον ένα άκρο της στο D. Άρα το D αποτελεί ένα κάλυμμα κορυφών του G. □

Αναγωγή του Καλύμματος Κορυφών στο Δέντρο Steiner. Έστω γράφημα $G = (V, E)$ στο οποίο ελέγχουμε αν υπάρχει κάλυμμα κορυφών C με $|C| \leq k$. Δημιουργούμε ένα πλήρες γράφημα G' με $|V| + |E|$ κορυφές, μία κορυφή u για κάθε $u \in V$, και μια κορυφή z_e για κάθε $e \in E$.

Στο G', μήκος 1 έχουν οι ακμές που συνδέουν είτε κορυφές u, v που αντιστοιχούν σε κορυφές του G, είτε μια κορυφή z_e, που αντιστοιχεί σε ακμή $e = \{u, v\}$ του G, με τα άκρα της u και v. Τα μήκος των υπόλοιπων ακμών του G' καθορίζεται από τις αποστάσεις των άκρων τους στο υπογράφημα του G' που περιλαμβάνει μόνο τις ακμές μήκους 1 (βλ. Σχήμα 4). Έτσι:

- Μήκος 2 έχουν οι ακμές του G' που συνδέουν είτε δύο κορυφές z_e και $z_{e'}$ για τις οποίες οι αντίστοιχες ακμές e και e' έχουν κοινό άκρο στο G, είτε δύο κορυφές u και z_e που η μια αντιστοιχεί σε κορυφή και η άλλη σε ακμή του G .
- Μήκος 3 έχουν οι ακμές του G' που συνδέουν δύο κορυφές z_e και $z_{e'}$ για τις οποίες οι αντίστοιχες ακμές e και e' δεν έχουν κοινό άκρο στο G.

Εξ' ορισμού, τα μήκη των ακμών ικανοποιούν την τριγωνική ανισότητα. Θέτουμε ακόμη $B = k + |E| - 1$. Στο στιγμιότυπο του Δέντρου Steiner ελέγχουμε αν υπάρχει συνεκτικό υπογράφημα του G' που περιλαμβάνει τις κορυφές του συνόλου $Z = \{z_e : e \in E\}$, δηλ. όλες τις κορυφές του G' που αντιστοιχούν σε ακμές του G, και έχει συνολικό μήκος μικρότερο ή ίσο του B.

Η κατασκευή του G' από το G και ο υπολογισμός του μήκους των ακμών μπορούν να γίνουν σε πολυωνυμικό χρόνο. Θα δείξουμε ότι το G έχει κάλυμμα κορυφών C με $|C| \leq k$ ανν στο G' υπάρχει δέντρο Steiner που καλύπτει τις κορυφές του Z και έχει συνολικό μήκος μικρότερο ή ίσο του $k + |E| - 1$.

Έστω κάλυμμα κορυφών C του G με $|C| \leq k$. Θεωρούμε το υπογράφημα του G' που περιλαμβάνει τις κορυφές του C, $k-1$ ακμές μήκους 1 που τις συνδέουν μεταξύ τους, τις κορυφές του Z, και για κάθε κορυφή $z_e \in Z$, την ακμή μήκους 1 που συνδέει την z_e με το άκρο της που ανήκει στο C (αν και τα δύο άκρα της e ανήκουν στο C, συνδέουμε την z_e με ένα από τα δύο άκρα). Το υπογράφημα αυτό είναι συνεκτικό γιατί το C "καλύπτει" όλες τις ακμές του G, περιλαμβάνει όλες τις κορυφές του Z, και έχει συνολικό βάρος $k + |E| - 1$.

Για το αντίστροφο, θεωρούμε συνδετικό υπογράφημα T του G' που περιλαμβάνει όλες τις κορυφές του Z και έχει συνολικό μήκος μικρότερο ή ίσο του $k + |E| - 1$. Το σημαντικό είναι πως μπορούμε να θεωρήσουμε ότι το T είναι δέντρο και περιλαμβάνει μόνο ακμές του G' με μήκος 1. Αν δεν ισχύει αυτό, μπορούμε να μετατρέψουμε το T σε ένα δέντρο T' που περιλαμβάνει όλες τις κορυφές του Z, έχει μόνο ακμές μήκους 1, και συνολικό μήκος μικρότερο ή ίσο από αυτό του T. Η μετατροπή μπορεί να γίνει ως εξής:

- Αν το T περιλαμβάνει ακμή $\{z_e, z_{e'}\}$ μήκους 3, όπου $e = \{u, v\}$ και $e' = \{u', v'\}$ ακμές στο G, την αντικαθιστούμε με τις ακμές $\{z_e, u\}$, $\{u, u'\}$, και $\{u', z_{e'}\}$, όλες μήκους 1.
- Αν το T περιλαμβάνει ακμή μήκους 2, αυτή:
  - Είτε θα είναι της μορφής $\{z_e, z_{e'}\}$, όπου οι αντίστοιχες ακμές $e = \{u, v\}$ και $e' = \{u, v'\}$ έχουν κοινό άκρο στο G (εδώ το u). Τότε αντικαθιστούμε την ακμή $\{z_e, z_{e'}\}$ με τις ακμές $\{z_e, u\}$ και $\{u, z_{e'}\}$, αμφότερες μήκους 1.
  - Είτε θα είναι της μορφής $\{z_e, w\}$, όπου $e = \{u, v\}$ ακμή και w κορυφή του G. Τότε αντικαθιστούμε την ακμή $\{z_e, w\}$ με τις ακμές $\{z_e, u\}$ και $\{u, w\}$, αμφότερες μήκους 1.

Αν το γράφημα που προκύπτει δεν είναι δέντρο, μπορούμε να το μετατρέψουμε σε δέντρο χωρίς να αυξήσουμε το βάρος του.

Εστιάζουμε λοιπόν στην περίπτωση που το T περιλαμβάνει μόνο ακμές μήκους 1. Αφού το συνολικό μήκος του T είναι μικρότερο ή ίσο του $k + |E| - 1$, το T περιλαμβάνει το πολύ $k + |E|$ κορυφές. Αφού το T περιλαμβάνει όλες τις κορυφές του Z (το πλήθος τους είναι $|E|$), το σύνολο C των κορυφών του T που αντιστοιχούν σε κορυφές του G αποτελείται από k κορυφές το πολύ. Επειδή το T είναι συνεκτικό και οι κορυφές του Z δεν συνδέονται μεταξύ τους με ακμές μήκους 1, κάθε κορυφή $z_e \in Z$ συνδέεται με ακμή μήκους 1 με κάποια κορυφή του C. Από τον ορισμό των μηκών, αυτή είναι κάποιο άκρο της ακμής e στο G. Δηλαδή, για κάθε ακμή e του G, τουλάχιστον ένα από τα άκρα της (αντιστοιχεί σε κορυφή που) ανήκει στο C. Άρα το C ορίζει ένα κάλυμμα κορυφών στο G με αριθμό κορυφών μικρότερο ή ίσο του k. $\square$

Αναγωγή του Κυρίαρχου Συνόλου στην Χωροθέτηση Υπηρεσιών. Έστω γράφημα $G = (V, E)$ στο οποίο ελέγχουμε αν υπάρχει κυρίαρχο σύνολο D με $|D| \leq k$. Δημιουργούμε ένα πλήρες γράφημα G' με σύνολο κορυφών V (ίδιο με αυτό του G). Το κόστος κάθε κορυφής του G' είναι ίσο με 2. Οι ακμές του G' που υφίστανται στο G έχουν μήκος 1, οι υπόλοιπες έχουν μήκος 2. Εξ΄ ορισμού τα μήκη των ακμών ικανοποιούν την τριγωνική ανισότητα. Θέτουμε ακόμη $B = |V| + k$.

Το G', μαζί με τα μήκη των ακμών, το κόστη των κορυφών, και το B αποτελούν ένα στιγμιότυπο για το πρόβλημα Χωροθέτησης Υπηρεσιών. Δεδομένου του G, το στιγμιότυπο αυτό υπολογίζεται σε πολυωνυμικό χρόνο. Θα δείξουμε ότι το G έχει κυρίαρχο σύνολο D με $|D| \leq k$ ανν υπάρχει λύση για την Χωροθέτηση Υπηρεσιών στο G' με συνολικό κόστος μικρότερο ή ίσο του $|V| + k$.

Πράγματι, κάθε κυρίαρχο σύνολο D, $|D| \leq k$, στο G ορίζει μια λύση για την χωροθέτηση υπηρεσιών στο G' με συνολικό κόστος μικρότερο ή ίσο του $|V| + k$. Το κόστος εγκατάστασης της υπηρεσίας στις κορυφές του D είναι $2|D|$, και το κόστος εξυπηρέτησης των κορυφών που δεν ανήκουν στο D είναι $|V| - |D|$, αφού όλες συνδέονται με ακμή μήκους 1 με κάποια κορυφή του D (επειδή το D είναι κυρίαρχο σύνολο στο G).

Αντίστροφα, έστω μια λύση για την Χωροθέτηση Υπηρεσιών στο G' η οποία ορίζεται από το σύνολο $C \subseteq V$ και έχει συνολικό κόστος μικρότερο ή ίσο του $|V|+k$. Χωρίς βλάβη της γενικότητας, θεωρούμε ότι κάθε κορυφή του $V \setminus C$ συνδέεται με ακμή μήκους 1 με κάποια κορυφή του C (αν υπάρχει $v \in V \setminus C$ που συνδέεται με ακμές μήκους 2 με όλες τις κορυφές του C, μπορούμε να προσθέσουμε την v στο C χωρίς να αυξήσουμε το συνολικό κόστος). Αφού το συνολικό κόστος είναι $2|C| + (|V| - |C|) \leq |V| + k$, έχουμε ότι $|C| \leq k$. Επιπλέον κάθε κορυφή που δεν ανήκει στο C συνδέεται (στο G) με ακμή με κάποια κορυφή του C. Συνεπώς το C αποτελεί ένα κυρίαρχο σύνολο του G με k κορυφές το πολύ. $\square$