Πανεπιστήμιο Πατρών

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογιάς Υπολογιστών

Τελική εργασία

Vertex Cover Problem

Συγγραφέας: Δημήτριος Δήμου

May 27, 2016

Contents

Chapter 1

Set cover problems

Το set cover problem είναι ένα κλασσικό πρόβλημα στον τομέα της συνδιαστικής βελτιστοποίησης και της θεωρίας υπολογιστών η μελέτη του οποίου έχει οδηγήση στην ανάπτυξη θεμελιώδων τεχνικών στο πεδίο των προσσεγιστικών αλγορίθμων. Λόγω της γενικής του διατύπωσης βρίσκει εφαρμογές σε μια εύρεια γμάμα προβλημάτων.

1.1 Διατύπωση

Το πρόβλημα μπορεί να διατυπωθεί με διάφορους τρόπους, είτε ως πρόβλημα βελτιστοποίησης όπου ζητείται ο ελάχιστος αριθμός υποσυνόλων ή αν έχει ανατεθεί συνάρτηση κόστους στα υποσύνολα ζητείται το σύνολο με το ελάχιστο κόστος είτε ως πρόβλημα απόφασης.

Απλή διατύπωση

Δεδομένου ενός σύμπαντος U αποτελούμενο από n στοιχεία, ενός συνόλου από υποσύνολα του U, $S = \{S_1, ..., S_k\}$ τέτοια ώστε η ένωσή τους να είναι το σύνολο U, βρες το ελάχιστο υποσύνολο του S που καλύπτει όλα τα στοιχεία του U.

Με συνάρτηση κόστους

Δεδομένου ενός σύμπαντος U αποτελούμενο από n στοιχεία, μιας συλλογής από υποσύνολα του U, $S = S_1, ..., S_k$, και μίας συνάρτησης κόστους $c : S \to \boldsymbol{Q}^+$, βρες το υποσύνολο του S με το ελάχιστο κόστος που καλύπτει όλα τα στοιχεία του U.

Πρόβλημα απόφασης

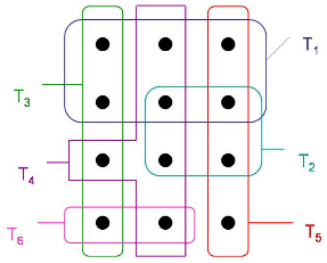
Δεδομένου ενός σύμπαντος U αποτελούμενο από n στοιχεία, μιας συλλογής από υποσύνολα του U, $S = S_1, ..., S_k$, και ενός ακεραίου k, αποφάσισε αν υπάρχει υποσύνολο του S με το πολύ k στοιχεία που καλύπτει όλα τα στοιχεία του U.

Παράδειγμα

Έστω το σύμπαν $U = \{1, 2, 3, 4, 5\}$ και η συλλογή από υποσύνολα του $S = \{\{1, 2, 3\}, \{2, 4\}, \{3, 4\}, \{4, 5\}\}$. Η ένωση των στοιχείων του S καλύπτει το U. Η ελάχιστη συλλογή υποσυνόλων του S που καλύπτει το U είναι τα : $\{\{1, 2, 3\}, \{4, 5\}\}$.

Το ελάχιστο set cover είναι το σύνολο $S' = \{T_3, T_4, T_5\}$

FIGURE 1.1: Set cover example



1.2 NP-completness

Το set cover decision problem είναι ένα από τα 21 NP-complete προβλήματα του Karp που αποδείχθηκε ότι είναι NP-complete το 1972. Αυτό σημαίνει ότι ανήκει στην κλάση NP δηλαδή δεδομένου ενός σύμπαντος U, μιας συλλγής S από υποσύνολα, ενός ακεραίου k και μίας λύσης S' η λύση αυτή μπορεί να επαληθευτεί σε πολυωνιμικό χρόνο όσον αφορά το μέγεθος των στοιχείων της εισόδου. Επίσης ανήκει και στην κλάση NP-hard. Αυτό οδήγησε στην ανάπτυξη προσσεγιστικών αλγορίθμων για την επίλυση του προβλήματος αυτού.

1.3 Λύσεις

1.3.1 Πρόβλημα ακέραιου προγραμματισμού

Το minimum set cover problem μπορεί να διατυπωθεί ως το ακόλουθο πρόβλημα ακέραιου προγραμματισμού

$$min\{\sum_{S\in\mathcal{S}} x_S\}$$

subject to

$$\sum_{S:e\in\mathcal{S}} x_S \geq 1, \quad \forall e \in \mathcal{U}$$

$$x_S \in \{0,1\}$$

Επειδή το πρόβλημα ακέραιου προγραμματισμού είναι NP-hard χαλαρώνουμε τους περιορισμούς του προβλήματος για το x_S και το ανάγουμε σε πρόβλημα γραμμικού περιορισμού το οποίο λύνεται σε πολυωνυμικό χρόνο. Έτσι καταλήγουμε στο πρόβλημα:

$$min\{\sum_{S\in\mathcal{S}} x_S\}$$

s.t.

$$\sum_{S:e\in\mathcal{S}} x_S \geq 1, \quad \forall e \in \mathcal{U}$$

$$x_S \in [0,1]$$

Επειδή το integrality gap αυτού του προβλήματος είναι το πολύ $\log n$ η χαλάρωση του δίνει factor-$\log n$ προσεγγιστικό αλγόριθμο.
Αν κάθε στοιχείο εμφανίζεται το πολύ σε $\mathcal{F}$ τότε μπορεί να βρεθεί λύση σε

πολυωνυμικό χρόνο η οποία προσεγγίζει το βέλτιστο με παράγοντα $\mathcal{F}$ χρησιμοποιόντας το πρόβλημα γραμμικού προγραμματισμού.

1.3.2 Άπληστος αλγόριθμος

Υπάρχει και ένας άπληστος αλγόριθμος για προσέγγιση σε πολυωνυμικό χρόνο ο οποίος διαλέγει τα σύνολα με έναν μόνο κανόνα: σε κάθε στάδιο διαλέγει το σύνολο που περιέχει τον μεγαλύτερο αριθμό από ακάλυπτα στοιχεία.

Αλγόριθμος:

1. $C \leftarrow 0$
2. While $C \neq \mathcal{U}$ do
   - (α') Find the set whose cost effectiveness is smallest, say S_i.
     Let $a = \frac{c(S_i)}{|S_i - C|}$.
     Pick S_i and $\forall e \in S_i - C$, set $price(e) = a$.
   - (β') $C \leftarrow S_i \cup C$
3. Output C

1.4 Hitting set

Το πρόβλημα set cover είναι ισοδύναμο με το πρόβλημα hitting set. Αν σε έναν διμερή γράφο το ένα σύνολο κόμβων $\mathcal{U}$ αντιπροσωπεύει τα υποσύνολα $\mathcal{S}$ του σύμπαντος, το άλλο σύνολο κόμβων $\mathcal{V}$ αντιπροσωπεύει τα στοιχεία του σύμπαντος και οι ακμές αντιπροσωπεύουν την συμπερίληψη ενός στοιχείου σε ένα σύνολο τότε βρίσκουμε των ελάχιστο αριθμό κόμβων του συνόλου $\mathcal{U}$ που καλύπτει όλους τους κόμβους του συνόλου $\mathcal{V}$.

1.5 Εφαρμογές

IBM finds computer viruses (wikipedia) elements- 5000 known viruses sets- 9000 substrings of 20 or more consecutive bytes from viruses, not found in 'good' code A set cover of 180 was found It suffices to search for these 180 substrings to verify the existence of known computer viruses. Another example: Consider General Motors needs to buy a certain amount of varied supplies and there are suppliers that offer various deals for different combina tions of materials (Supplier A: 2 tons of steel + 500 tiles for $x; Supplier B: 1 ton of steel + 2000 tiles for $y; etc.). You could use set covering to find the best way to get all the materials while minimizing cost

Chapter 2

Vertex cover problem

Το vertex cover ενός γράφου είναι ένα σύνολο κόμβων τέτοιο ώστε κάθε ακμή του γράφου είναι προσκύμενη σε τουλάχιστον ένα κομβό του συνόλου αυτού. Το πρόβλημα εύρεσης του ελάχιστου vertex cover είναι κλασσικό προβλημα στον τομέα της συνδιαστικής βελτιστοποίησης και της θεωρίας υπολογιστών και κλασσικό παράδειγμα NP-hard προβλήματος βελτιστοποίησης.

2.1 Διατύπωση

Το vertex cover V' ενός μη κατευθυντικού γράφου $G = (V, E)$ ειναι ένα υποσύνολο του V τέτοιο ώστε:

$$\forall uv \in E \Rightarrow u \in V' \vee v \in V'$$

Ένα τέτοι σύνολο λέμε οτί καλύπτει τις ακμές του G. Το ελάχιστο vertex cover ενός γράφου G είναι το σύνολο V' με τον μικρότερο αριθμό στοιχείων.

FIGURE 2.1: Vertex cover

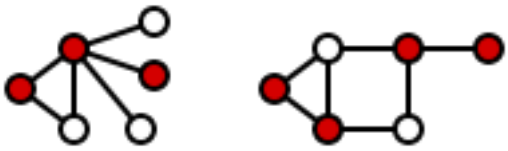
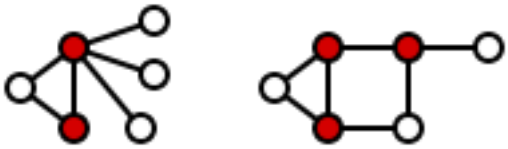


FIGURE 2.2: Minimum vertex cover



2.2 NP-completness

2.3 Λύσεις

2.3.1 Πρόβλημα ακέραιου προγραμματισμού

Το minimum vertex cover problem μπορεί να διατυπωθεί ως το ακόλουθο πρόβλημα ακέραιου προγραμματισμού

$c : S \to \boldsymbol{Q}^+$

$$min\{\sum_{u \in V} c(v)x_v\}$$

subject to

$$x_u + x_v \geq 1, \quad \forall (u,v) \in E$$

$$x_S \in \{0,1\}$$

Επειδή το πρόβλημα ακέραιου προγραμματισμού είναι NP-hard χαλαρώνουμε τους περιορισμούς του προβλήματος για το x_v και το ανάγουμε σε πρόβλημα γραμμικού περιορισμού το οποίο λύνεται σε πολυωνυμικό χρόνο. Έτσι καταλήγουμε στο πρόβλημα:

$$min\{\sum_{S \in \mathcal{S}} x_S\}$$

subject to

$$\sum_{S:e \in \mathcal{S}} x_S \geq 1, \quad \forall e \in \mathcal{U}$$

$$x_v \geq 0, \quad v \in V$$

Επειδή το integrality gap αυτού του προβλήματος είναι το πολύ $\log n$ η χαλάρωση του δίνει factor-$\log n$ προσεγγιστικό αλγόριθμο.
Αν κάθε στοιχείο εμφανίζεται το πολύ σε $\mathcal{F}$ τότε μπορεί να βρεθεί λύση σε πολυωνυμικό χρόνο η οποία προσεγγίζει το βέλτιστο με παράγοντα $\mathcal{F}$ χρησιμοποιόντας το πρόβλημα γραμμικού προγραμματισμού.

2.3.2 Άπληστος αλγόριθμος

Αλγόριθμος:

1. $C \leftarrow 0$
2. While $C \neq \mathcal{U}$ do
   - (α′) Find the set whose cost effectiveness is smallest, say S_i.
     Let $a = \frac{c(S_i)}{|S_i - C|}$.
     Pick S_i and $\forall e \in S_i - C$, set $price(e) = a$.
   - (β′) $C \leftarrow S_i \cup C$
3. Output C

2.4 Εφαρμογές