

---

---

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ  
Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών  
και Πληροφορικής

Αποδοτικοί αλγόριθμοι για την κατανομή  
ενέργειας σε ασύρματα δίκτυα

Διδακτορική Διατριβή

ΣΤΑΥΡΟΣ ΑΘΑΝΑΣΟΠΟΥΛΟΣ

Επιβλέπων: Χρήστος Κακλαμάνης, Καθηγητής

Πάτρα, Ιούνιος 2009

---

---



# Αποδοτικοί αλγόριθμοι για την κατανομή ενέργειας σε ασύρματα δίκτυα

Διδακτορική Διατριβή

ΣΤΑΥΡΟΣ ΑΘΑΝΑΣΟΠΟΥΛΟΣ

Επαμελής Εξεταστική Επιτροπή :

Χρήστος Κακλαμάνης, Καθηγητής (Επιβλέπων)  
Χρήστος Ζαρολιάγκης, Καθηγητής  
Παύλος Σπυράκης, Καθηγητής  
Εμμανουήλ Βαρβαρίγος, Καθηγητής  
Ιωάννης Καραγιάννης, Λέκτορας  
Σταύρος Κοσμαδάκης, Καθηγητής  
Σωτήρης Νικολετσέας, Επίκουρος Καθηγητής



Η εκπόνηση της παρούσας Διδακτορικής Διατριβής υποστηρίχθηκε μερικώς από τη Γενική Γραμματεία Έρευνας και Τεχνολογίας στα πλαίσια του έργου 03ΕΔ599 “Θεωρητική Ανάλυση και Πειραματική Μελέτη Αλγορίθμων για Ενεργειακά Αποδοτική Επικοινωνία σε Ασύρματα Δίκτυα Τύπου Ad Hoc” (ΠΕΝΕΔ 2003).



# Αποδοτικοί αλγόριθμοι για την κατανομή ενέργειας σε ασύρματα δίκτυα

ΣΤΑΥΡΟΣ ΑΘΑΝΑΣΟΠΟΥΛΟΣ

## Περίληψη

Στην παρούσα διδακτορική διατριβή, ασχολούμαστε με ζητήματα που ανακύπτουν σε ασύρματα δίκτυα επικοινωνίας, δηλ. δίκτυα που βασίζονται σε τηλεπικοινωνιακή υποδομή όπως τα κυψελικά δίκτυα κινητής τηλεφωνίας, δίκτυα αυτόνομων ασύρματων εκπομπών όπως τα ασύρματα δίκτυα τύπου ad hoc, κτλ. Τα ασύρματα δίκτυα επικοινωνίας διαφόρων τύπων έχουν εξελιχθεί σημαντικά τα τελευταία χρόνια. Ειδικότερα, τα ασύρματα αδόμητα δίκτυα (ή αλλιώς ασύρματα δίκτυα τύπου ad hoc) έχουν προσελκύσει το έντονο ενδιαφέρον της επιστημονικής κοινότητας λόγω των πολλών εφαρμογών που έχουν κυρίως σε περιπτώσεις όπου δεν είναι δυνατή ή επιθυμητή η ολική ή μερική κάλυψη μέσω υποδομής με βάση την ενσύρματη δικτύωση (π.χ., επικοινωνία σε δυσπρόσιτες ή απομακρυσμένες περιοχές, φυσικές καταστροφές, στρατιωτικές εφαρμογές, κλπ.).

Όπως και στα παραδοσιακά ενσύρματα δίκτυα, σημαντικό πρόβλημα αποτελεί η εγκαθίδρυση σχημάτων επικοινωνίας όπως διάδοση (broadcasting, multicasting), επικοινωνία όλων με όλους (gossiping, all-to-all communication), και επικοινωνία σε ομάδες (group communication). Για την επικοινωνία απαιτείται η κατανάλωση ενέργειας στους κόμβους του δικτύου και, λαμβάνοντας υπόψη ότι τα αδόμητα ασύρματα δίκτυα χρησιμοποιούν κόμβους με περιορισμένα αποθέματα ενέργειας, είναι απαραίτητη η ορθολογιστική χρήση αυτής της ενέργειας κατά την επικοινωνία. Αυτό μπορεί να σημαίνει ότι είναι επιθυμητή είτε η ελαχιστοποίηση της συνολικής ενέργειας που καταναλώνεται στους κόμβους του δικτύου για επικοινωνία ή η ελαχιστοποίηση της μέγιστης ενέργειας ώστε να επιτυγχάνεται όσο το δυνατό μεγαλύτερος χρόνος ζωής όλων των κόμβων του δικτύου. Στη διατριβή εξετάζουμε αλγόριθμους για την εγκαθίδρυση διαφορετικών σχημάτων επικοινωνίας σε αδόμητα ασύρματα δίκτυα όπου βασικό κριτήριο για την εκτίμηση της απόδοσής τους θα είναι η κατανάλωση ενέργειας που επιφέρουν στο δίκτυο. Μοντελοποιούμε τα δίκτυα με ειδικά γραφήματα και τα αντίστοιχα προβλήματα επικοινωνίας σαν προβλήματα συνδυαστικής βελτιστοποίησης στα γραφήματα αυτά.

Τα αποτελέσματά μας περιλαμβάνουν νέους αλγόριθμους που βελτιώνουν προηγούμενα γνωστά σχετικά αποτελέσματα και νέα κάτω φράγματα. Με κεντρικό στόχο την αποδοτική κατανομή ενέργειας σε ασύρματα δίκτυα, η μελέτη μας έχει διπλό χαρακτήρα: από τη μια πλευρά, ασχολούμαστε με μελέτη και ανάλυση θεμελιωδών προβλημάτων της Θεωρητικής Επιστήμης των Υπολογιστών (όπως, π.χ., το πρόβλημα Κάλυψης με Σύνολα). Τέτοια προβλήματα, και ειδικές περιπτώσεις τους, παρουσιάζουν εξαιρετικό ενδιαφέρον αφού χρησιμοποιούνται (μεταξύ άλλων) συχνά για τη μοντελοποίηση προβλημάτων ενεργειακά αποδοτικής επικοινωνίας σε ασύρματα δίκτυα. Επιπλέον, προτείνουμε και αναλύουμε νέους αλγόριθμους για συγκεκριμένα σενάρια επικοινωνίας σε σύγχρονα ασύρματα δίκτυα. Από την άλλη πλευρά, μελετάμε και εκτιμούμε πειραματικά την απόδοση αρκετών αλγορίθμων και τεχνικών (από τη βιβλιογραφία αλλά και νέων) για ενεργειακά αποδοτική επικοινωνία σε ασύρματα δίκτυα.





# Efficient algorithms for energy allocation in wireless networks

STAVROS ATHANASSOPOULOS

## Abstract

In this dissertation, we study issues arising in wireless communication networks, i.e., networks based on telecommunication infrastructure like cellular wireless networks, networks of autonomous wireless transmitters like ad hoc wireless networks, and so on. Wireless networks have received significant attention during the recent years. Especially, ad hoc wireless networks - for which unlike traditional wired networks or cellular wireless networks, no wired backbone infrastructure is installed - emerged due to their potential applications in emergency disaster relief, battlefield, etc.

Like in traditional wired networks, an important problem concerns the establishment of communication patterns like broadcasting, multicasting, gossiping, all-to-all communication, and group communication. Communication then requires energy consumption at network nodes, and given that in ad hoc wireless networks energy is a scarce resource, it is of paramount importance to use it efficiently when establishing communication patterns. In such a setting, it is usually pursued that either the total energy consumed at networks nodes or the maximum energy consumed at any network node is minimized so that the network lifetime is prolonged as long as possible. Herein, we present and analyze - theoretically and experimentally - algorithms for guaranteeing the establishment of various communication patterns in ad hoc wireless networks and evaluate their performance in terms of their energy-efficiency. We represent these networks using graphs and model the corresponding communication problems as combinatorial optimization problems in such graphs.

Our results include new algorithms which improve previously known relevant results as well as new lower bounds. Our main objective being the efficient energy allocation in wireless networks, our study is of dual

character: on the one hand, we study and analyze fundamental problems of Theoretical Computer Science (like, e.g., Set Cover); such problems, as well as special cases of them, are highly interesting since they usually model energy-efficient communication problems in wireless networks. Furthermore, we propose and analyse new algorithms for particular communication scenaria in modern wireless networks. On the other hand, we experimentally study and evaluate several algorithms and techniques (both from the literature and new ones) for energy-efficient communication in wireless networks.



# Περιεχόμενα

|   |           |
|---|-----------|
| <b>1 Συνεισφορά της διατριβής</b>   | <b>1</b>  |
| <b>2 Προβλήματα κάλυψης με σύνολα</b>   | <b>5</b>  |
| 2.1 Εισαγωγή . . . . .  | 5         |
| 2.2 Άπληστοι αλγόριθμοι . . . . .   | 7         |
| 2.3 Περιγραφή των νέων αλγορίθμων . . . . .   | 9         |
| 2.4 Ανάλυση μέσω γραμμικών προγραμμάτων . . . . .   | 14        |
| 2.5 Απόδειξη του Θεωρήματος 5 . . . . .   | 19        |
| 2.6 Απόδειξη του Θεωρήματος 6 . . . . .   | 26        |
| 2.7 Επεκτάσεις . . . . .  | 34        |
| <b>3 Εύρεση γεννητικού δάσους αστέρων και εξοικονόμηση χρωμάτων</b>                                 | <b>35</b> |
| 3.1 Εισαγωγή . . . . .  | 36        |
| 3.2 Περιγραφή του αλγορίθμου . . . . .  | 40        |
| 3.3 Ανάλυση . . . . .   | 42        |
| 3.3.1 Ένα παραμετροποιημένο γραμμικό πρόγραμμα . . . . .  | 42        |
| 3.3.2 Απόδειξη του φράγματος για το λόγο προσέγγισης . . . . .                                      | 46        |
| 3.4 Εξοικονόμηση χρωμάτων . . . . .   | 48        |
| 3.5 Ένα κάτω φράγμα για αλγόριθμους τοπικής αναζήτησης . . . . .                                    | 49        |
| <b>4 Ενεργειακά αποδοτική επικοινωνία σε ασύρματα δίκτυα με πολλαπλά μέσα ασύρματης διασύνδεσης</b> | <b>53</b> |
| 4.1 Εισαγωγή . . . . .  | 54        |
| 4.2 Ορισμοί και συμβολισμοί . . . . .   | 57        |
| 4.3 Ένας αλγόριθμος για τη συμμετρική περίπτωση . . . . .   | 58        |
| 4.4 Ομαδική επικοινωνία στη συμμετρική περίπτωση . . . . .  | 61        |
| 4.5 Ομαδική επικοινωνία στη μη συμμετρική περίπτωση . . . . .                                       | 62        |

|          |  |           |
|----------|--|-----------|
| <b>5</b> | <b>Ενεργειακά αποδοτική διάδοση μηνυμάτων σε ασύρματα αδόμη-<br/>τα δίκτυα</b> | <b>67</b> |
| 5.1      | Εισαγωγή . . . . .   | 67        |
| 5.2      | Αλγόριθμοι τοπικής αναζήτησης . . . . .  | 72        |
| 5.3      | Επαυξητικοί αλγόριθμοι . . . . .   | 74        |
| 5.4      | Πειραματικά αποτελέσματα . . . . .   | 80        |
| <b>6</b> | <b>Επίλογος</b>  | <b>89</b> |

# Κεφάλαιο 1

## Συνεισφορά της διατριβής

Στα πλαίσια της παρούσας διδακτορικής διατριβής με τίτλο “Αποδοτικοί αλγόριθμοι για την κατανομή ενέργειας σε ασύρματα δίκτυα”, ασχολούμαστε με ασύρματα αδόμητα δίκτυα ή αλλιώς ασύρματα δίκτυα τύπου *ad hoc*, τα οποία τις τελευταίες δεκαετίες έχουν αναπτυχθεί ραγδαία λόγω των ενδεχόμενων εφαρμογών τους σε στρατιωτικές επιχειρήσεις, στην αντιμετώπιση έκτακτων καταστροφών, κτλ. [79].

Στα δίκτυα αυτά, σε αντίθεση με τα παραδοσιακά κυβελικά δίκτυα, οι κόμβοι, δηλαδή οι συσκευές που απαρτίζουν το δίκτυο, όπως κινητά τηλέφωνα, φορητοί υπολογιστές, αισθητήρες, κτλ., επικοινωνούν μεταξύ τους ασύρματα και δεν υπάρχει ενσύρματη δικτυακή υποδομή για την υποστήριξη της ασύρματης επικοινωνίας. Επιπλέον, οι κόμβοι διαθέτουν κάποιου είδους μπαταρία, δηλαδή περιορισμένη ενέργεια και αποτελούν μέρος του δικτύου όσο η ενέργειά τους δεν έχει εξαντληθεί. Επομένως, τα δίκτυα αυτά δεν έχουν συγκεκριμένη δομή για αυτό λέγονται και αδόμητα ή *ad hoc*. Η διατήρηση της ενέργειας των κόμβων σε τέτοια δίκτυα είναι εξαιρετικά σημαντική για την παράταση της διάρκειας ζωής των κόμβων και κατά συνέπεια και του ολόκληρου του δικτύου.

Στα πλαίσια αυτού του σεναρίου ενδιαφέρουσες ερευνητικά προκλήσεις σχετικά με την αποδοτική κατανάλωση της ενέργειας που διαθέτουν οι κόμβοι αφορούν αφενός στο σχεδιασμό, ανάπτυξη, θεωρητική ανάλυση και πειραματική αξιολόγηση αλγορίθμων οι οποίοι δεδομένων των εγγενών χαρακτηριστικών των δικτύων αυτών εξασφαλίζουν συγκεκριμένα σχήματα επικοινωνίας με χαμηλή κατανάλωση ενέργειας και αφετέρου στη μελέτη θεμελιωδών προβλημάτων της θεωρίας γραφημάτων: τέτοια προβλήματα, και ειδικές περιπτώσεις τους, παρουσιάζουν εξαιρετικό ενδιαφέρον αφού χρησιμοποιούν-

ται (μεταξύ άλλων) συχνά για τη μοντελοποίηση προβλημάτων ενεργειακά αποδοτικής επικοινωνίας σε ασύρματα δίκτυα, τη διερεύνηση της υπολογιστικής τους δυσκολίας καθώς και των περιορισμών στην επίτευξη βέλτιστων ως προς την κατανάλωση ενέργειας λύσεων. Μερικά από τα πιο χαρακτηριστικά θεμελιώδη γραφοθεωρητικά προβλήματα που έχουν χρησιμοποιηθεί συχνά στη βιβλιογραφία για να συλλάβουν τα ιδιαίτερα χαρακτηριστικά προβλημάτων επικοινωνίας σε ασύρματα δίκτυα είναι το πρόβλημα Κάλυψης με Σύνολα (Set cover), το πρόβλημα εύρεσης Ελάχιστου Γεννητικού δένδρου (Minimum Spanning Tree), το πρόβλημα εύρεσης δασών ή/και δένδρων Steiner (Steiner Forest ή Steiner Tree), το πρόβλημα εύρεσης Συνόλου Κυριαρχίας (Dominating Set), το πρόβλημα εύρεσης Ανεξάρτητου Συνόλου (Independent Set) και πολλά άλλα.

Η συνεισφορά της παρούσας διδακτορικής διατριβής αναφέρεται και στις δύο κατευθύνσεις που περιγράφονται παραπάνω, και αφορά στη μελέτη θεμελιωδών προβλημάτων της θεωρίας γραφημάτων και παρουσίαση νέων βελτιωμένων τεχνικών και αποτελεσμάτων, στην ανάπτυξη νέων, αποδεδειγμένα αποδοτικών ως προς την κατανάλωση ενέργειας αλγορίθμων επικοινωνίας σε ασύρματα αδόμητα δίκτυα αλλά και στην πειραματική σύγκριση και αξιολόγηση αλγορίθμων για συγκεκριμένα δημοφιλή σχήματα επικοινωνίας σε τέτοια δίκτυα.

Ειδικότερα, η συνεισφορά μας είναι συνοπτικά η εξής:

- Ασχοληθήκαμε με το θεμελιώδες γραφοθεωρητικό πρόβλημα Κάλυψης με Σύνολα (Set Cover) και προτείναμε μια νέα τεχνική ανάλυσης συνδυαστικών αλγορίθμων με χρήση γραμμικών προγραμμάτων [4] την οποία στη συνέχεια εφαρμόσαμε για το πρόβλημα εύρεσης μέγιστου γεννητικού δάσους αστέρων (spanning star forest) [6].

Συγκεκριμένα, στο Κεφάλαιο 2, μελετάμε το πρόβλημα κάλυψης με σύνολα και ενδιαφέρουσες παραλλαγές του. Παρουσιάζουμε νέους συνδυαστικούς προσεγγιστικούς αλγόριθμους για το πρόβλημα  $k$ -κάλυψης συνόλων. Προηγούμενες προσεγγίσεις έχουν βασισθεί σε επεκτάσεις του άπληστου αλγόριθμου μέσω αποδοτικού χειρισμού μικρών συνόλων. Οι νέοι αλγόριθμοι επεκτείνουν περαιτέρω τις προηγούμενες προσεγγίσεις χρησιμοποιώντας την ιδέα του υπολογισμού μεγάλων ομάδων στοιχείων και στη συνέχεια της ομαδοποίησής τους σε σύνολα μεγάλου μεγέθους. Τα αποτελέσματά μας βελτιώνουν τα καλύτερα γνωστά φράγματα προσέγγισης για το πρόβλημα  $k$ -κάλυψης συνόλων για κάθε τιμή



του  $k \geq 6$ . Η τεχνική που χρησιμοποιούμε για την ανάλυση παρουσιάζει επιπλέον ανεξάρτητα ενδιαφέρον· το πάνω φράγμα για τον παράγοντα προσέγγισης επιτυγχάνεται φράσσοντας την αντικειμενική τιμή ενός γραμμικού προγράμματος η οποία ‘αποκαλύπτει’ το λόγο προσέγγισης του υπό εξέταση αλγορίθμου (*factor-revealing linear program*).

Στο Κεφάλαιο 3, παρουσιάζουμε έναν απλό αλγόριθμο για το πρόβλημα εύρεσης μέγιστου δάσους γεννητικού αστέρα. Λαμβάνουμε υπόψη το γεγονός ότι το πρόβλημα αποτελεί ειδική περίπτωση του συμπληρωματικού προβλήματος κάλυψης συνόλου και προσαρμόζουμε έναν αλγόριθμο των Duh και Fürer για την επίλυσή του. Αποδεικνύουμε ότι ο αλγόριθμος αυτός υπολογίζει  $193/240 \approx 0.804$ -προσεγγιστικά δάση γεννητικών αστέρων. Το αποτέλεσμα αυτό βελτιώνει ένα προηγούμενο άνω φράγμα με τιμή 0.71 των Chen και άλλων. Αν και ο αλγόριθμος είναι καθαρά συνδυαστικός, η ανάλυσή μας ορίζει ένα γραμμικό πρόγραμμα που χρησιμοποιεί μια παράμετρο  $f$  το οποίο είναι επιλύσιμο για τιμές της παραμέτρου  $f$  που δεν είναι μικρότερες από το λόγο προσέγγισης του αλγορίθμου. Η ανάλυση είναι αυστηρή και, το ενδιαφέρον είναι, ότι μπορεί να εφαρμοστεί και σε συμπληρωματικές εκδοχές του προβλήματος κάλυψης συνόλου όπως η εξοικονόμηση χρωμάτων. Δίνει την ίδια εγγύηση προσέγγισης με τιμή  $193/240$  που οριακά βελτιώνει το προηγούμενο γνωστό κάτω φράγμα των Duh και Fürer. Αποδεικνύουμε επίσης ότι, γενικά, μια φυσική κλάση αλγορίθμων τοπικής αναζήτησης δε δίνουν καλύτερα από  $1/2$ -προσεγγιστικά δάση γεννητικών αστέρων.

- Παρουσιάσαμε νέους αλγόριθμους για διατήρηση συνεκτικότητας και για ομαδική επικοινωνία στην περίπτωση που οι κόμβοι του δικτύου υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης [7].

Στο Κεφάλαιο 4, μελετάμε προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης. Σε τέτοια δίκτυα, δύο κόμβοι μπορούν να επικοινωνήσουν αν είναι αρκετά κοντά και διαθέτουν κάποιο κοινό μέσο ασύρματης διασύνδεσης. Η ενεργοποίηση ενός μέσου ασύρματης διασύνδεσης επιφέρει ένα κόστος που αντανakλά την ενέργεια που καταναλώνεται όταν κάποιος κόμβος χρησιμοποιεί το μέσο αυτό. Διακρίνουμε μεταξύ της συμμετρικής και της μη συμμετρικής περίπτωσης, με βάση το κόστος ενεργοποίησης για κάθε ασύρματο μέσο διασύνδεσης είναι το ίδιο για όλους τους κόμβους ή όχι. Για τη συμμετρική περίπτωση, παρουσιάζουμε έναν  $(3/2 + \epsilon)$ -

προσεγγιστικό αλγόριθμο για το πρόβλημα πλήρους διασύνδεσης με ελάχιστο κόστος ενεργοποίησης, βελτιώνοντας ένα προηγούμενο φράγμα με τιμή 2. Για τη μη συμμετρική περίπτωση, αποδεικνύουμε ότι το πρόβλημα διασύνδεσης δεν είναι προσεγγίσιμο στα πλαίσια ενός παράγοντα υπολογαριθμικού ως προς το πλήθος των κόμβων και παρουσιάζουμε ένα λογαριθμικό προσεγγιστικό αλγόριθμο για μια γενικότερη περίπτωση που μοντελοποιεί την ομαδική επικοινωνία.

- Επίσης, μελετήσαμε πειραματικά αλγόριθμους για τον υπολογισμό αποδοτικών ως προς την ενέργεια δένδρων μετάδοσης (multicasting) σε ασύρματα αδόμητα δίκτυα, από τη βιβλιογραφία αλλά και νέους, με σαφώς βελτιωμένη απόδοση [6].

Στο Κεφάλαιο 5, μελετάμε αλγόριθμους για τον υπολογισμό αποδοτικών ως προς την ενέργεια δένδρων μετάδοσης (multicasting) σε ασύρματα αδόμητα δίκτυα. Τέτοιοι αλγόριθμοι είτε ξεκινούν από μια κενή λύση η οποία σταδιακά επαυξάνεται για να δώσει ένα δένδρο μετάδοσης (*επαυξητικοί αλγόριθμοι - augmentation algorithms*) είτε λαμβάνουν σαν είσοδο ένα αρχικό δένδρο μετάδοσης και εκτελούν 'περιπάτους' σε διαφορετικά δένδρα μετάδοσης για πεπερασμένο αριθμό βημάτων μέχρι να επιτευχθεί κάποια αποδεκτή μείωση στην κατανάλωση της ενέργειας (*αλγόριθμοι τοπικής αναζήτησης - local search algorithms*). Εστιάζουμε τόσο σε επαυξητικούς αλγόριθμους όσο και σε αλγόριθμους τοπικής αναζήτησης και συγκεκριμένα έχουμε υλοποιήσει αρκετούς υπάρχοντες αλγόριθμους από τη βιβλιογραφία αλλά και νέους. Συγκρίνουμε πειραματικά τους αλγόριθμους αυτούς σε τυχαία γεωμετρικά στιγμιότυπα του προβλήματος και επιτυγχάνουμε αποτελέσματα όσον αφορά στην αποδοτικότητα ως προς την ενέργεια των λύσεων που λαμβάνουμε. Παρουσιάζουμε επίσης αποτελέσματα σχετικά με το χρόνο εκτέλεσης των υλοποιήσεών μας. Επίσης διερευνούμε το κατά πόσον οι λύσεις που λαμβάνουμε από επαυξητικούς αλγόριθμους μπορούν να βελτιωθούν μέσω αλγορίθμων τοπικής αναζήτησης. Τα αποτελέσματά μας αποδεικνύουν ότι ένας από τους νέους αλγόριθμους που προτείνουμε και οι εκδοχές του επιτυγχάνουν τις πιο αποδοτικές ενεργειακά λύσεις και μάλιστα πολύ γρήγορα. Η έρευνά μας υποδεικνύει ιδιότητες γεωμετρικών στιγμιότυπων του προβλήματος που συντελούν στη βελτιωμένη απόδοση των επαυξητικών αλγορίθμων.

## Κεφάλαιο 2

# Προβλήματα κάλυψης με σύνολα

Το πρόβλημα κάλυψης με σύνολα αποτελεί θεμελιώδες πρόβλημα συνδυαστικής βελτιστοποίησης με πολλές εφαρμογές. Στιγμιότυπα του προβλήματος εμπεριέχουν ένα σύνολο στοιχείων  $V$  και μια συλλογή  $S$  υποσυνόλων του  $V$  ενώ στόχος είναι η επιλογή ενός υποσυνόλου του  $S$  με ελάχιστο πλήθος στοιχείων έτσι ώστε κάθε στοιχείο να καλύπτεται, δηλ., να περιέχεται σε τουλάχιστον ένα επιλεγμένο σύνολο.

### 2.1 Εισαγωγή

Το πρόβλημα κάλυψης με σύνολα και οι παραλλαγές του καθώς και συναφή με αυτό γραφοθεωρητικά προβλήματα, όπως το πρόβλημα κάλυψης κορυφών, το πρόβλημα εύρεσης συνόλου κυριαρχίας κτλ., σχετίζονται άμεσα με διάφορα 'πρακτικά' προβλήματα σχετικά με ενεργειακά αποδοτική επικοινωνία σε ασύρματα αδόμητα δίκτυα (ασύρματα δίκτυα τύπου ad hoc).

Οι Cardei και Du [18] (και επίσης η εργασία [60]) μελετούν το πρόβλημα κάλυψης σημείων στο οποίο πρέπει να παρακολουθείται ένα περιορισμένο πλήθος σημείων (που καλούνται στόχοι) με γνωστές θέσεις. Ένα μεγάλο πλήθος αισθητήρων διασπείρονται τυχαία σε μικρή απόσταση από τους στόχους και αποστέλλουν την πληροφορία που συλλέγουν σε έναν κεντρικό κόμβο. Κάθε στόχος πρέπει να παρακολουθείται συνεχώς από τουλάχιστον έναν αισθητήρα, υποθέτοντας ότι κάθε αισθητήρας μπορεί να παρακολουθεί όλους τους στόχους εντός της εμβέλειάς του. Η οργάνωση των αισθητήρων στο μέγιστο δυνατό αριθμό συνόλων που καλύπτουν την περιοχή δίνει τη δυνατότητα ενεργοποίησης των συνόλων αυτών διαδοχικά οπότε εξοικονομείται ενέργεια

και παρατείνεται η διάρκεια ζωής του δικτύου: μόνο οι αισθητήρες του τρέχοντος ενεργοποιημένου συνόλου επιτηρούν τους στόχους, ενώ όλοι οι υπόλοιποι είναι ανενεργοί. Ελαττώνοντας το χρονικό διάστημα κατά το οποίο ένας αισθητήρας είναι ενεργοποιημένος, αυξάνεται ο συνολικός χρόνος μέχρι η ενέργεια όλων των αισθητήρων να εξαντληθεί, και η διάρκεια της εφαρμογής αυξάνεται ανάλογα κατά έναν παράγοντα ίσο με τον αριθμό των ξένων μεταξύ τους συνόλων.

Όταν υλοποιείται ένα μεγάλο και πυκνό δίκτυο τυχαία τοποθετημένων αισθητήρων για την παρακολούθηση μιας περιοχής, το κατά πόσον καλύπτεται η περιοχή προσεγγίζεται από τη συνολική εμβέλεια των αισθητήρων. Δηλαδή, καλύπτοντας την περιοχή εμβέλειας κάθε αισθητήρα προσεγγίζουμε την κάλυψη κάθε σημείου στη συγκεκριμένη περιοχή. Μια μέθοδος που διασφαλίζει κάλυψη και συνεκτικότητα είναι ο σχεδιασμός του συνόλου των ενεργοποιημένων αισθητήρων σαν ένα συνεκτικό σύνολο κυριαρχίας. Ένα πρωτόκολλο για την κατασκευή ενός τέτοιου συνόλου κυριαρχίας παρουσιάζεται από τους Wu και Li με χρήση της τεχνικής μαρκαρίσματος (marking) [83].

Ο Wu και άλλοι στη εργασία [82] μελετούν επίσης το πρόβλημα της κάλυψης μέσω ενεργειακά αποδοτικών συνόλων κυριαρχίας. Γενικά, οι κόμβοι στο συνεκτικό σύνολο κυριαρχίας καταναλώνουν περισσότερη ενέργεια προκειμένου να χειριστούν ποικίλη κίνηση που περνά από αυτούς σε σχέση με κόμβους εκτός του συνόλου κυριαρχίας. Για να παραταθεί η διάρκεια ζωής κάθε κόμβου, και επομένως και του δικτύου μέσω εξισορρόπησης της συνολικής κατανάλωσης ενέργειας σε αυτό, οι κόμβοι θα πρέπει να επιλέγονται εναλλακτικά να αποτελούν συνεκτικό σύνολο κυριαρχίας. Προτείνεται έτσι μια μέθοδος επιλογής κόμβων κάλυψης με βάση τα ενεργειακά τους επίπεδα η οποία αποδεικνύεται πειραματικά [82] ότι έχει βελτιωμένη απόδοση σε σχέση με άλλες προϋπάρχουσες όσον αφορά στη διάρκεια ζωής του δικτύου. Το πρόβλημα κάλυψης κόμβων σχετίζεται στενά και με το πρόβλημα της διάδοσης (broadcasting), στο οποίο επιλέγεται ένας μικρός αριθμός κόμβων προώθησης [81]. Η επιλογή του συνόλου προώθησης για το πρόβλημα της διάδοσης είναι ανάλογη με το πρόβλημα κάλυψης σημείων, αφού και στα δύο στόχος είναι η εύρεση ενός μικρού συνόλου κάλυψης. Αξίζει να σημειωθεί ότι η τεχνική αυτή χρησιμοποιείται και στην περίπτωση κατευθυνόμενης διάχυσης μέσω διάδοσης [50]. Καθώς η πληροφορία προωθείται στο δίκτυο, οι κόμβοι αισθητήρες εγκαθιδρύουν αντίστροφες συνδέσεις επικοινωνίας με την πηγή με αποκεντρωμένο τρόπο, με τη διαφορά ότι τώρα όλοι οι αισθητήρες προωθούν τα δεδομένα.

## 2.2 Άπληστοι αλγόριθμοι

Ο φυσικός άπληστος αλγόριθμος για το πρόβλημα κάλυψης με σύνολα ξεκινάει με μια λύση χωρίς στοιχεία και την αυξάνει μέχρι όλα τα στοιχεία να καλυφθούν επιλέγοντας ένα σύνολο που περιέχει το μέγιστο αριθμό στοιχείων που δεν περιέχονται σε κανένα από τα σύνολα που έχουν ήδη επιλεγεί. Συμβολίζοντας με  $n$  τον αριθμό των στοιχείων, είναι γνωστό από τις αρχικές εργασίες των Johnson [52], Lovasz [63], και Chvátal [24] ότι ο άπληστος αλγόριθμος επιτυγχάνει λόγο προσέγγισης  $H_n$ . Η αυστηρότερη ανάλυση του Slavík [74] βελτιώνει το άνω φράγμα για το λόγο προσέγγισης σε  $\ln n - \ln \ln n + O(1)$ . Ασυμπτωτικά, τα φράγματα αυτά είναι αυστηρά χάρη στο γνωστό σημαντικό αποτέλεσμα μη προσεγγισιμότητας του Feige [38] το οποίο δηλώνει ότι, για κάθε σταθερό  $\epsilon > 0$ , δεν υπάρχει  $(1 - \epsilon) \ln n$ -προσεγγιστικός αλγόριθμος για το πρόβλημα κάλυψης με σύνολα με την υπόθεση ότι δεν υπάρχουν για όλα τα προβλήματα της κλάσης  $NP$  αλγόριθμοι με υποεκθετικούς χρόνους εκτέλεσης  $n^{O(\log \log n)}$ .

Μια ενδιαφέρουσα παραλλαγή είναι το πρόβλημα κάλυψης με σύνολα με μέγεθος το πολύ  $k$ . Χωρίς απώλεια της γενικότητας, μπορούμε να υποθέσουμε ότι στο  $S$  ισχύει η κλειστότητα ως προς τα υποσύνολα. Στην περίπτωση αυτή, ο άπληστος αλγόριθμος μπορεί ισοδύναμα να διατυπωθεί ως εξής:

**Άπληστες φάσεις:** Για  $i = k$  έως 1 κάνε:

Επίλεξε μέγιστη συλλογή από ξένα μεταξύ τους  $i$ -σύνολα.

Ένα  $i$ -σύνολο είναι ένα σύνολο που περιέχει ακριβώς  $i$  στοιχεία που δεν έχουν ήδη καλυφθεί και μια συλλογή  $T$  από ξένα μεταξύ τους  $i$ -σύνολα καλείται μέγιστη αν κάθε άλλο  $i$ -σύνολο τέμνει (δηλ. έχει κοινά στοιχεία με) κάποια από τα σύνολα του  $T$ .

Ένα αυστηρό φράγμα με τιμή  $H_k$  για το λόγο προσέγγισης του άπληστου αλγόριθμου είναι γνωστό για την περίπτωση αυτή [52, 63]. Επειδή το πρόβλημα έχει πληθώρα εφαρμογών για συγκεκριμένες τιμές του  $k$  και λόγω του ότι είναι ενδιαφέρον από τη σκοπιά της θεωρίας πολυπλοκότητας, ο σχεδιασμός αλγορίθμων με βελτιωμένους όρους δεύτερης τάξης στο λόγο προσέγγισής τους έχει τύχει ιδιαίτερης προσοχής. Μέχρι τώρα, υπάρχουν αλγόριθμοι με λόγο προσέγγισης  $H_k - c$  όπου  $c$  είναι μια θετική σταθερά. Οι Goldschmidt και άλλοι [40] ήταν οι πρώτοι που παρουσίασαν μια τροποποιημένη έκδοση του άπληστου αλγόριθμου με  $c = 1/6$ . Η τιμή αυτή βελτιώθηκε

σε  $1/3$  από το Halldórsson [44] και σε  $1/2$  από τους Duh και Fürer [34]. Πρόσφατα, ο Levin [59] βελτίωσε περαιτέρω τη σταθερά σε  $98/195 \approx 0.5026$  για  $k \geq 4$ . Στην αρνητική πλευρά, ο Trevisan [76] απέδειξε ότι, αν δεν υπάρχουν ντετερμινιστικοί αλγόριθμοι υποεκθετικού χρόνου για την κλάση  $NP$ , κανένας αλγόριθμος πολυωνυμικού χρόνου δεν έχει λόγο προσέγγισης  $\ln k - \Omega(\ln \ln k)$ .

Η βασική ιδέα που έχει χρησιμοποιηθεί για να βελτιωθεί η απόδοση του άπληστου αλγόριθμου είναι ο αποδοτικός χειρισμός μικρών συνόλων. Ο αλγόριθμος των Goldschmidt και άλλων [40] χρησιμοποιεί υπολογισμό ταιριασμάτων για να συμπεριλάβει όσο το δυνατόν περισσότερα στοιχεία σε σύνολα μεγέθους 2 όταν κανένα σύνολο μεγέθους τουλάχιστον 3 δεν περιέχει νέα στοιχεία. Οι αλγόριθμοι των Halldórsson [43, 44] και Duh και Fürer [34] χειρίζονται αποδοτικά σύνολα μεγέθους 3. Ο αλγόριθμος στην εργασία [34] βασίζεται σε μια ημι-τοπική τεχνική βελτιστοποίησης. Η βελτίωση του Levin [59] επεκτείνει τον αλγόριθμο της εργασίας [34] με αποδοτικό χειρισμό συνόλων μεγέθους 4.

Μια φυσική αλλά πλήρως αποδοτική ιδέα είναι η αντικατάσταση των φάσεων του άπληστου αλγόριθμου που σχετίζονται με μεγάλα σύνολα με φάσεις ομαδοποίησης συνόλων οι οποίες επίσης αποσκοπούν στη μεγιστοποίηση του αριθμού νέων στοιχείων που καλύπτονται από μεγάλα μέγιστα σύνολα. Ο παράγοντας προσέγγισης δε χειροτερεύει (λόγω τη συνθήκης μεγιστοποίησης) αλλά έχει παραμείνει ανοιχτό πρόβλημα στην εργασία [59] αν αυτό οδηγήσει σε κάποια βελτίωση στο φράγμα προσέγγισης. Στη συνέχεια, δείχνουμε ότι αντικαθιστώντας τις άπληστες φάσεις στους αλγόριθμους των Duh και Fürer με φάσεις ομαδοποίησης, επιτυγχάνουμε βελτιωμένα φράγματα προσέγγισης για κάθε  $k \geq 6$  που προσεγγίζουν την τιμή  $H_k - 0.5902$  για μεγάλες τιμές του  $k$ . Τα αποτελέσματα που παρουσιάζουμε στη συνέχεια εμφανίστηκαν στην εργασία [4].

Συγκεκριμένα, θα χρησιμοποιήσουμε αλγόριθμους για το πρόβλημα ομαδοποίησης συνόλων με μέγεθος το πολύ  $k$  που ορίζεται ως εξής. Ένα στιγμιότυπο του προβλήματος ομαδοποίησης συνόλων με μέγεθος το πολύ  $k$  περιλαμβάνει ένα σύνολο στοιχείων  $V$  και μια συλλογή  $S$  υποσυνόλων του  $V$  καθένα από τα οποία περιέχει ακριβώς  $k$  στοιχεία. Στόχος είναι η επιλογή όσο το δυνατόν περισσότερων ξένων μεταξύ τους συνόλων του  $S$ . Όταν  $k = 2$ , το πρόβλημα είναι ισοδύναμο με το πρόβλημα εύρεσης ταιριάσματος μέγιστου μεγέθους σε γραφήματα, και επομένως, είναι επιλύσιμο σε πολυωνυμικό χρόνο. Για  $k \geq 3$ , το πρόβλημα είναι APX-δύσκολο [53]: το καλύτερο γνωστό φράγμα μη προσεγγισιμότητας για μεγάλα  $k$  δηλώνει ότι δεν είναι δυνατόν να

προσεγγιστεί το πρόβλημα ομαδοποίησης συνόλων μεγέθους εντός  $\Omega\left(\frac{\log k}{k}\right)$  εκτός εάν  $P = NP$  [47]. Σημειώνουμε ότι κάθε μέγιστη συλλογή ξένων μεταξύ τους υποσυνόλων δίνει μια  $1/k$ -προσεγγιστική λύση. Οι καλύτεροι γνωστοί αλγόριθμοι έχουν λόγο προσέγγισης  $\frac{2-\epsilon}{k}$  για κάθε  $\epsilon > 0$  [48] και βασίζονται σε τοπική αναζήτηση· αυτοί είναι οι αλγόριθμοι που χρησιμοποιούνται κατά τις φάσεις ομαδοποίησης των δικών μας αλγορίθμων.

Η ανάλυσή μας βασίζεται στην ιδέα των γραμμικών προγραμμάτων (LP) των οποίων η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του υπό εξέταση αλγορίθμου η οποία εισήχθη αρχικά σε ένα διαφορετικό πλαίσιο στην εργασία [51] για την ανάλυση προσεγγιστικών αλγορίθμων για το πρόβλημα τοποθέτησης μέσων/πόρων (facility location). Δείχνουμε ότι ο όρος προσέγγισης φράσσεται άνω από τη μέγιστη αντικειμενική τιμή ενός τέτοιου γραμμικού προγράμματος. Επομένως, δεν απαιτείται ρητή τεκμηρίωση για τη δομή της λύσης που υπολογίζεται από τους αλγόριθμους. Αντίθετα, οι εγγυήσεις απόδοσης των διάφορων φάσεων χρησιμοποιούνται σε μαύρα κουτιά κατά τον ορισμό των περιορισμών του LP.

Στη συνέχεια, παρουσιάζουμε τους αλγόριθμους των Duh και Furer [34] καθώς και τις δικές μας τροποποιήσεις στην Ενότητα 2.3. Η τεχνική της ανάλυσης παρατίθεται στην Ενότητα 2.4 όπου παρουσιάζουμε τα λήμματα για το γραμμικό πρόγραμμα του οποίου η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του αλγορίθμου και τις αποδείξεις των βασικών αποτελεσμάτων μας. Παρουσιάζουμε τα συμπεράσματά μας στην Ενότητα 2.7.

## 2.3 Περιγραφή των νέων αλγορίθμων

Στην ενότητα αυτή δίνουμε αρχικά μια συνοπτική περιγραφή των σχετικών αποτελεσμάτων της εργασίας [34]· και στη συνέχεια, παρουσιάζουμε τους δικούς μας αλγόριθμους και τα βασικά μας αποτελέσματα.

Εκτός από τον άπληστο αλγόριθμο, αλγόριθμοι τοπικής αναζήτησης έχουν χρησιμοποιηθεί για το πρόβλημα κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$ , ειδικότερα για μικρές τιμές του  $k$ . Η καθαρή τοπική αναζήτηση ξεκινάει από αυθαίρετο σύνολο κάλυψης και εξελίσσεται σε βήματα. Σε κάθε βήμα, η τρέχουσα λύση βελτιώνεται με αντικατάσταση ενός σταθερού αριθμού συνόλων από έναν (ευελπιστούμε μικρότερο) αριθμό άλλων συνόλων έτσι ώστε να επιτευχθεί ένα νέο σύνολο κάλυψης. Οι Duh και Furer εισήγαγαν την τεχνική της ημι-τοπικής βελτιστοποίησης η οποία επεκτείνει την καθαρή τοπική αναζήτηση. Στα πλαίσια του προβλήματος κάλυψης 3-συνόλων, η

βασική ιδέα πίσω από την ημι-τοπική βελτιστοποίηση είναι ότι από τη στιγμή που επιλεχθούν τα σύνολα μεγέθους 3, ο υπολογισμός του ελάχιστου αριθμού συνόλων μεγέθους 2 και 1 για να ολοκληρωθεί το σύνολο κάλυψης μπορεί να πραγματοποιηθεί σε πολυωνυμικό χρόνο μέσω του υπολογισμού ενός ταιριάσματος. Επομένως, ένα ημι-τοπικό βήμα  $(s, t)$ -βελτίωσης για το πρόβλημα κάλυψης 3-συνόλων αποτελείται από τη διαγραφή μέχρι το πολύ  $t$  3-συνόλων από το τρέχον σύνολο κάλυψης και την εισαγωγή μέχρι το πολύ  $s$  3-συνόλων καθώς και των ελάχιστων απαραίτητων 2-συνόλων και 1-συνόλων που ολοκληρώνουν το σύνολο κάλυψης. Η ποιότητα του συνόλου κάλυψης που παράγεται από μια βελτίωση ορίζεται από το συνολικό αριθμό συνόλων στο σύνολο κάλυψης ενώ στην περίπτωση δύο συνόλων κάλυψης ίδιου μεγέθους, προτιμάται αυτό με το μικρότερο αριθμό 1-συνόλων. Ημι-τοπική βελτιστοποίηση για  $k \geq 4$  είναι εν πολλοίς ίδια· οι τοπικές βελτιώσεις ορίζονται τώρα με βάση σύνολα μεγέθους τουλάχιστον 3 ενώ επιλέγονται βέλτιστα 2-σύνολα και 1-σύνολα. Η ανάλυση στην εργασία [34] δείχνει ότι η βέλτιστη επιλογή των παραμέτρων  $(s, t)$  είναι  $(2, 1)$  καθώς και ότι, κάπως αναπάντεχα, μεγαλύτερες τιμές των  $s$  και  $t$  δε δίνουν γενικά καλύτερες λύσεις.

**Θεώρημα 1** (Duh και Furer [34]). *Θεωρούμε ένα στιγμιότυπο του προβλήματος κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$  του οποίου η βέλτιστη λύση περιέχει  $a_i$   $i$ -σύνολα. Τότε, ο ημι-τοπικός αλγόριθμος  $(2, 1)$ -βελτιστοποίησης έχει κόστος το πολύ  $a_1 + a_2 + \sum_{i=3}^k \frac{i+1}{3} a_i$ .*

*Απόδειξη. (συνοπτικά)* Η απόδειξη στην εργασία [34] προχωράει ως εξής. Έστω  $b_i$  ο αριθμός των  $i$ -συνόλων στη λύση. Αρχικά παρατηρούμε ότι  $\sum_{i=1}^k i b_i = \sum_{i=1}^k i a_i$ . Μετά, αποδεικνύονται οι παρακάτω δύο ιδιότητες της ημι-τοπικής  $(2, 1)$ -βελτιστοποίησης:  $b_1 \leq a_1$  και  $b_1 + b_2 \leq \sum_{i=1}^k a_i$ . Το θεώρημα προκύπτει από τις τρεις ανισότητες.  $\square$

Ο αλγόριθμος αυτός έχει αποτελέσει τη βάση για τους ακόλουθους αλγόριθμους που προσεγγίζουν το πρόβλημα κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$ . Τους καλούμε  $\text{GSLI}_{k,\ell}$  και  $\text{GRSLI}_{k,\ell}$ , αντίστοιχα.

*Αλγόριθμος  $\text{GSLI}_{k,\ell}$ .*

**Άπληστες φάσεις:** Για  $i = k$  έως  $\ell + 1$  κάνε:

Επίλεξε μια μέγιστη συλλογή  $i$ -συνόλων.

**Ημι-τοπική φάση βελτιστοποίησης:** Θέτουμε  $i = \ell$ . Εκτελούμε τον ημι-τοπικό αλγόριθμο  $(2, 1)$ -βελτιστοποίησης στο υπόλοιπο στιγμιότυπο.



**Θεώρημα 2** (Duh και Fürer [34]). *Ο αλγόριθμος  $\text{GSLI}_{k,4}$  έχει λόγο προσέγγισης  $H_k - 5/12$ .*

*Ο αλγόριθμος  $\text{GRSLI}_{k,\ell}$*

**Άπληστες φάσεις:** Για  $i = k$  έως  $\ell + 1$  κάνε:

Επίλεξε μια μέγιστη συλλογή  $i$ -συνόλων.

**Περιορισμένες φάσεις:** Για  $i = \ell$  έως 4 κάνε:

Επίλεξε μια μέγιστη συλλογή  $i$ -συνόλων έτσι ώστε η επιλογή αυτών των  $i$ -συνόλων να μην αυξάνει τον αριθμό των 1-συνόλων στην τελική λύση.

**Ημι-τοπική φάση βελτιστοποίησης:** Θέτουμε  $i = 3$ . Εκτελούμε τον ημι-τοπικό αλγόριθμο βελτιστοποίησης στο υπόλοιπο στιγμιότυπο.

**Θεώρημα 3** (Duh και Fürer [34]). *Ο αλγόριθμος  $\text{GRSLI}_{k,5}$  έχει λόγο προσέγγισης  $H_k - 1/2$ .*

Θα τροποποιήσουμε τους παραπάνω αλγόριθμους αντικαθιστώντας κάθε άπληστη φάση με μία φάση ομαδοποίησης για το χειρισμό συνόλων όχι πολύ μικρού μεγέθους. Χρησιμοποιούμε τους αλγόριθμους τοπικής αναζήτησης των Hurkens και Schrijver [48] σε κάθε φάση ομαδοποίησης. Οι τροποποιημένοι αλγόριθμοι καλούνται  $\text{PSLI}_{k,\ell}$  και  $\text{PRSLI}_{k,\ell}$ , αντίστοιχα.

Ένας αλγόριθμος τοπικής αναζήτησης για ομαδοποίηση συνόλων χρησιμοποιεί μια σταθερή παράμετρο  $p$  (άτυπα, αυτό αποτελεί ένα άνω φράγμα για τον αριθμό των τοπικών βελτιώσεων που πραγματοποιούνται σε κάθε βήμα λοσαλ) και, ξεκινώντας από μία άδεια ομαδοποίηση  $\Pi$ , επαναλαμβανόμενα ενημερώνει την  $\Pi$  αντικαθιστώντας κάθε σύνολο με  $s < p$  σύνολα της  $\Pi$  με  $s + 1$  σύνολα έτσι ώστε να διατηρείται η ορθότητα και μέχρι να μην είναι δυνατή άλλη αντικατάσταση. Προφνώς, ο αλγόριθμος έχει πολυωνυμικό χρόνο εκτέλεσης. Η ανάλυσή του παρουσιάζεται στην εργασία [48]. (Δείτε επίσης την εργασία [43] για σχετικά ζητήματα).

**Θεώρημα 4** (Hurkens και Schrijver [48]). *Ο αλγόριθμος τοπικής αναζήτησης για  $t$ -ομαδοποίηση συνόλων που εκτελεί το πολύ  $p$  τοπικές βελτιώσεις σε κάθε βήμα έχει λόγο προσέγγισης  $\rho_t \geq \frac{2(t-1)^{r-t}}{t(t-1)^{r-t}}$  αν  $p = 2r - 1$ , και  $\rho_t \geq \frac{2(t-1)^{r-2}}{t(t-1)^{r-2}}$  αν  $p = 2r$ .*

Σα συμπέρασμα, για κάθε σταθερό  $\epsilon > 0$ , επιτυγχάνουμε έναν  $\frac{2-\epsilon}{t}$ -προσεγγιστικό αλγόριθμο για  $t$ -ομαδοποίηση συνόλων χρησιμοποιώντας  $p = O(\log_t 1/\epsilon)$  τοπικές βελτιώσεις. Οι αλγόριθμοί μας  $\text{PSL}_{k,\ell}$  και  $\text{PRSL}_{k,\ell}$  απλά αντικαθιστούν κάθε άπληστη φάση των αλγορίθμων  $\text{GSL}_{k,\ell}$  και  $\text{GRSL}_{k,\ell}$ , αντίστοιχα, με την παρακάτω φάση ομαδοποίησης:

**Φάσεις ομαδοποίησης:** Για  $i = k$  έως  $\ell + 1$  κάνε:

Επίλεξε μια μέγιστη συλλογή ξένων μεταξύ τους  $i$ -συνόλων χρησιμοποιώντας έναν  $\frac{2-\epsilon}{i}$ -προσεγγιστικό αλγόριθμο τοπικής αναζήτησης για  $i$ -ομαδοποίηση συνόλων.

Το πρώτο βασικό μας αποτέλεσμα (Θεώρημα 5) αναφέρεται στην απόδοση του αλγορίθμου  $\text{PSL}_{k,\ell}$  (για  $\ell = 4$  που αποτελεί και την καλύτερη επιλογή).

**Θεώρημα 5.** Για κάθε σταθερό  $\epsilon > 0$ , ο αλγόριθμος  $\text{PSL}_{k,4}$  έχει λόγο προσέγγισης το πολύ  $H_{k/2} + \frac{1}{6} + \epsilon$  για άρτιο  $k \geq 6$ , το πολύ  $2H_k - H_{\frac{k-1}{2}} - \frac{11}{9} + \epsilon$  για  $k \in \{5, 7, 9, 11, 13\}$ , και το πολύ  $H_{\frac{k-1}{2}} + \frac{1}{6} + \frac{2}{k} - \frac{1}{k-1} + \epsilon$  για περιττά  $k \geq 15$ .

Ο αλγόριθμος  $\text{PSL}_{k,4}$  επιτυγχάνει καλύτερη απόδοση από τον  $\text{GSL}_{k,4}$  για  $k \geq 5$ , από τον αλγόριθμο  $\text{GRSL}_{k,5}$  για  $k \geq 19$ , καθώς και από τη βελτίωση του Levin [59] για  $k \geq 21$ . Για μεγάλες τιμές του  $k$ , το φράγμα προσέγγισης τείνει στην τιμή  $H_k - c$  με  $c = \ln 2 - 1/6 \approx 0.5264$ . Ο αλγόριθμος  $\text{PSL}_{k,\ell}$  αποτελεί το κύριο θέμα στη συνέχεια για να εισάγουμε την τεχνική της ανάλυσης. Όπως θα δούμε στην επόμενη ενότητα, το γραμμικό πρόγραμμα του οποίου η αντικειμενική τιμή 'αποκαλύπτει' το λόγο προσέγγισης του αλγορίθμου είναι απλούστερο σε αυτή την περίπτωση. Ο αλγόριθμος  $\text{PRSL}_{k,\ell}$  είναι ακόμα καλύτερος· η απόδοσή του (για  $\ell = 5$ ) παρουσιάζεται στο επόμενο θεώρημα.

**Θεώρημα 6.** Για κάθε σταθερό  $\epsilon > 0$ , ο αλγόριθμος  $\text{PRSL}_{k,5}$  έχει λόγο προσέγγισης το πολύ  $2H_k - H_{\frac{k-1}{2}} - \frac{77}{60} + \epsilon$  για περιττά  $k \geq 7$ , το πολύ  $\frac{461}{240} + \epsilon$  για  $k = 6$ , και το πολύ  $2H_k - H_{k/2} - \frac{77}{60} + \frac{2}{k} - \frac{1}{k-1} + \epsilon$  για άρτια  $k \geq 8$ .

Ο αλγόριθμος  $\text{PRSL}_{k,5}$  πετυχαίνει καλύτερο λόγο προσέγγισης από τον αλγόριθμο του Levin [59] για κάθε  $k \geq 6$ . Για παράδειγμα, ο λόγος προσέγγισης  $\frac{461}{240} \approx 1.9208$  για το πρόβλημα κάλυψης 6-συνόλων βελτιώνει το προηγούμενο φράγμα με τιμή  $H_6 - \frac{98}{195} \approx 1.9474$ . Για μεγάλες τιμές του  $k$ , το φράγμα προσέγγισης τείνει στην τιμή  $H_k - c$  με  $c = 77/60 - \ln 2 \approx 0.5902$ . Στον Πίνακα 2.1 παρουσιάζεται μια σύγκριση των αλγορίθμων στους οποίους αναφερόμαστε σε αυτή την ενότητα.

Πίνακας 2.1: Σύγκριση του λόγου προσέγγισης των αλγορίθμων  $\text{GSLI}_{k,4}$ ,  $\text{PSLI}_{k,4}$ ,  $\text{GRSLI}_{k,5}$ , του αλγορίθμου της εργασίας [59] και του αλγορίθμου  $\text{PRSLI}_{k,5}$  για αρκετές τιμές του  $k$ .

| $k$        | $\text{GSLI}_{k,4}$ [34] | $\text{PSLI}_{k,4}$ | $\text{GRSLI}_{k,5}$ [34] | [59]           | $\text{PRSLI}_{k,5}$ |
|------------|--------------------------|---------------------|---------------------------|----------------|----------------------|
| 3          | 1.3333                   | 1.3333              | 1.3333                    | 1.3333         | 1.3333               |
| 4          | 1.6667                   | 1.6667              | 1.5833                    | 1.5808         | 1.5833               |
| 5          | 1.8667                   | 1.8444              | 1.7833                    | 1.7801         | 1.7833               |
| 6          | 2.0333                   | 2                   | 1.95                      | 1.9474         | 1.9208               |
| 7          | 2.1762                   | 2.1429              | 2.0929                    | 2.0903         | 2.0690               |
| 8          | 2.3012                   | 2.25                | 2.2179                    | 2.2153         | 2.1762               |
| 9          | 2.4123                   | 2.3524              | 2.3290                    | 2.3264         | 2.2917               |
| 10         | 2.5123                   | 2.45                | 2.4290                    | 2.4264         | 2.3802               |
| 19         | 3.1311                   | 3.0453              | 3.0477                    | 3.0452         | 2.9832               |
| 20         | 3.1811                   | 3.0956              | 3.0977                    | 3.0952         | 3.0305               |
| 21         | 3.2287                   | 3.1409              | 3.1454                    | 3.1428         | 3.0784               |
| 22         | 3.2741                   | 3.1865              | 3.1908                    | 3.1882         | 3.1217               |
| 50         | 4.0825                   | 3.9826              | 3.9992                    | 3.9966         | 3.9187               |
| 75         | 4.4847                   | 4.3814              | 4.4014                    | 4.3988         | 4.3178               |
| 100        | 4.7707                   | 4.6659              | 4.6874                    | 4.6848         | 4.6021               |
| μεγάλο $k$ | $H_k - 0.4167$           | $H_k - 0.5264$      | $H_k - 0.5$               | $H_k - 0.5026$ | $H_k - 0.5902$       |

## 2.4 Ανάλυση μέσω γραμμικών προγραμμάτων

Οι αποδείξεις μας για την εγγύηση προσέγγισης των αλγορίθμων μας προκύπτουν ουσιαστικά μέσω του υπολογισμού άνω φραγμάτων για την αντικειμενική τιμή γραμμικών προγραμμάτων η οποία ‘αποκαλύπτει’ το λόγο προσέγγισης του υπό εξέταση αλγορίθμου. Οι περιορισμοί αυτών των γραμμικών προγραμμάτων συλλαμβάνουν απλές σταθερές που διατηρούνται στις φάσεις του αλγορίθμου. Στη συνέχεια, περιγράφουμε πώς επιτυγχάνουμε αυτά τα γραμμικά προγράμματα.

Κάθε διαφορετική φάση των αλγορίθμων μπορεί να είναι μια φάση ομαδοποίησης, μια περιορισμένη φάση, ή η τελική ημι-τοπική φάση βελτιστοποίησης και προσδιορίζεται από την τιμή της μεταβλητής  $i$ : ο όρος ‘φάση που σχετίζεται με  $i$ ’ χρησιμοποιείται για να γίνει αναφορά σε κάθε διαφορετική φάση.

Θεωρούμε ένα στιγμιότυπο  $(V, S)$  του προβλήματος κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$ . Για κάθε φάση των αλγορίθμων που σχετίζεται με  $i$  ( $i = \ell, \dots, k$  για τον αλγόριθμο  $\text{PSLI}_{k,\ell}$  και  $i = 3, \dots, k$  για τον αλγόριθμο  $\text{PRSLI}_{k,\ell}$ ), θεωρούμε το στιγμιότυπο  $(V_i, S_i)$  όπου  $V_i$  περιέχει τα στοιχεία του  $V$  που δεν έχουν καλυφθεί σε προηγούμενες φάσεις και το  $S_i$  περιέχει τα σύνολα του  $S$  που περιέχουν μόνο στοιχεία του  $V_i$ . Συμβολίζουμε με  $\text{OPT}^i$  μια βέλτιστη λύση για το στιγμιότυπο  $(V_i, S_i)$ : επίσης, συμβολίζουμε τη βέλτιστη λύση  $\text{OPT}^k$  για το  $(V_k, S_k) = (V, S)$  με  $\text{OPT}$ . Αφού στο  $S$  ισχύει η κλειστότητα ως προς τα υποσύνολα, χωρίς απώλεια της γενικότητας, μπορούμε να υποθέσουμε ότι το  $\text{OPT}^i$  περιέχει ξένα μεταξύ τους σύνολα. Επιπλέον, είναι σαφές ότι  $|\text{OPT}^{i-1}| \leq |\text{OPT}^i|$  για  $i \leq k$ , δηλ.,  $|\text{OPT}^i| \leq |\text{OPT}|$ .

Για μια φάση του αλγορίθμου  $\text{PSLI}_{k,\ell}$  ή  $\text{PRSLI}_{k,\ell}$  που σχετίζεται με  $i$ , συμβολίζουμε με  $a_{i,j}$  το λόγο του αριθμού των  $j$ -συνόλων στο  $\text{OPT}^i$  προς τον αριθμό  $|\text{OPT}|$  των συνόλων στη βέλτιστη λύση. Αφού  $|\text{OPT}^i| \leq |\text{OPT}|$ , έχουμε ότι

$$\sum_{j=1}^i a_{i,j} \leq 1. \quad (2.1)$$

Ο αλγόριθμος για  $i$ -ομαδοποίηση συνόλων που εκτελείται κατά τη φάση ομαδοποίησης που σχετίζεται με  $i$  προσθέτει στα  $i$ -σύνολα τα στοιχεία του  $V_i \setminus V_{i-1}$ . Επειδή  $V_{i-1} \subseteq V_i$ , ο αριθμός τους είναι

$$|V_i \setminus V_{i-1}| = |V_i| - |V_{i-1}| = \left( \sum_{j=1}^i j a_{i,j} - \sum_{j=1}^{i-1} j a_{i-1,j} \right) |\text{OPT}|. \quad (2.2)$$

Συμβολίζουμε με  $\rho_i$  το λόγο προσέγγισης του αλγορίθμου για την  $i$ -ομαδοποίηση συνόλων που εκτελείται κατά τη φάση που σχετίζεται με  $i$ . Αφού στην έναρξη της φάσης ομαδοποίησης που σχετίζεται με  $i$ , υπάρχουν τουλάχιστον  $a_{i,i}|OPT|$   $i$ -σύνολα, ο αλγόριθμος για  $i$ -ομαδοποίηση συνόλων υπολογίζει τουλάχιστον  $\rho_i a_{i,i}|OPT|$   $i$ -σύνολα, δηλ., καλύπτει τουλάχιστον  $i\rho_i a_{i,i}|OPT|$  στοιχεία συνόλων του  $OPT^i$ . Επομένως,  $|V_i \setminus V_{i-1}| \geq i\rho_i a_{i,i}|OPT|$ , και (2.2) δίνει

$$\sum_{j=1}^{i-1} ja_{i-1,j} - \sum_{j=1}^{i-1} ja_{i,j} - i(1 - \rho_i)a_{i,i} \leq 0. \quad (2.3)$$

Μέχρι τώρα, έχουμε ορίσει όλους τους περιορισμούς για το γραμμικό πρόγραμμα του οποίου η αντικειμενική τιμή 'αποκαλύπτει' το λόγο προσέγγισης του αλγορίθμου  $PSL_{k,\ell}$ . Στη συνέχεια, φράσσουμε άνω τον αριθμό των συνόλων που υπολογίζει ο αλγόριθμος  $PSL_{k,\ell}$  ως εξής. Έστω  $t_i$  ο αριθμός των  $i$ -συνόλων που υπολογίζονται από τον αλγόριθμο για  $i$ -ομαδοποίηση συνόλων που εκτελείται κατά τη φάση ομαδοποίησης που σχετίζεται με  $i \geq \ell + 1$ . Προφανώς,

$$t_i = \frac{1}{i}|V_i \setminus V_{i-1}| = \left( \frac{1}{i} \sum_{j=1}^i ja_{i,j} - \frac{1}{i} \sum_{j=1}^{i-1} ja_{i-1,j} \right) |OPT|. \quad (2.4)$$

Από το Θεώρημα 1, έχουμε ότι

$$t_\ell \leq \left( a_{\ell,1} + a_{\ell,2} + \sum_{j=3}^{\ell} \frac{j+1}{3} a_{\ell,j} \right) |OPT|. \quad (2.5)$$

Επομένως, από (2.4) και (2.5), προκύπτει ότι η εγγύηση προσέγγισης του αλγορίθμου  $PSL_{k,\ell}$  είναι

$$\begin{aligned} \frac{\sum_{i=\ell+1}^k t_i}{|OPT|} &\leq \sum_{i=\ell+1}^k \frac{1}{i} \left( \sum_{j=1}^i ja_{i,j} - \sum_{j=1}^{i-1} ja_{i-1,j} \right) + a_{\ell,1} + a_{\ell,2} + \sum_{j=3}^{\ell} \frac{j+1}{3} a_{\ell,j} \\ &= \frac{1}{k} \sum_{j=1}^k ja_{k,j} + \sum_{i=\ell+1}^{k-1} \frac{1}{i(i+1)} \sum_{j=1}^i ja_{i,j} + \frac{\ell}{\ell+1} a_{\ell,1} \\ &\quad + \frac{\ell-1}{\ell+1} a_{\ell,2} + \sum_{j=3}^{\ell} \left( \frac{j+1}{3} - \frac{j}{\ell+1} \right) a_{\ell,j}. \end{aligned} \quad (2.6)$$

Άρα, ένα άνω φράγμα για το λόγο προσέγγισης του αλγορίθμου  $\text{PSLI}_{k,\ell}$  προκύπτει αν μεγιστοποιήσουμε το δεξί μέρος της (2.6) υπό τους περιορισμούς (2.1) για  $i = \ell, \dots, k$  και (2.3) για  $i = \ell + 1, \dots, k$  με μεταβλητές  $a_{i,j} \geq 0$  για  $i = \ell, \dots, k$  και  $j = 1, \dots, i$ . Τυπικά, έχουμε αποδείξει την ακόλουθη δήλωση.

**Λήμμα 7.** *Ο λόγος προσέγγισης του αλγορίθμου  $\text{PSLI}_{k,\ell}$  όταν χρησιμοποιείται ένας  $\rho_i$ -προσεγγιστικός αλγόριθμος για  $i$ -ομαδοποίηση συνόλων κατά τη φάση  $i$  για  $i = \ell + 1, \dots, k$  φράσσεται άνω από τη μέγιστη αντικειμενική τιμή του ακόλουθου γραμμικού προγράμματος:*

$$\begin{aligned}
& \text{maximize} && \frac{1}{k} \sum_{j=1}^k j a_{k,j} + \sum_{i=\ell+1}^{k-1} \frac{1}{i(i+1)} \sum_{j=1}^i j a_{i,j} + \frac{\ell}{\ell+1} a_{\ell,1} \\
& && + \frac{\ell-1}{\ell+1} a_{\ell,2} + \sum_{j=3}^{\ell} \left( \frac{j+1}{3} - \frac{j}{\ell+1} \right) a_{\ell,j} \\
& \text{subject to} && \sum_{j=1}^i a_{i,j} \leq 1, i = \ell, \dots, k \\
& && \sum_{j=1}^{i-1} j a_{i-1,j} - \sum_{j=1}^{i-1} j a_{i,j} - i(1 - \rho_i) a_{i,i} \leq 0, i = \ell + 1, \dots, k \\
& && a_{i,j} \geq 0, i = \ell, \dots, k, j = 1, \dots, i
\end{aligned}$$

Κάθε φάση ομαδοποίησης ή κάθε περιορισμένη φάση του αλγορίθμου  $\text{PRSLI}_{k,\ell}$  ικανοποιεί τους περιορισμούς (2.3): μια περιορισμένη φάση που σχετίζεται με  $i = 4, \dots, \ell$  υπολογίζει μια μέγιστη  $i$ -ομαδοποίηση συνόλων και, επομένως,  $\rho_i = 1/i$  σε αυτή την περίπτωση.

Επιπλέον, οι περιορισμένες φάσεις θέτουν κι άλλους περιορισμούς. Συμβολίζουμε με  $b_1$ ,  $b_2$ , και  $b_3$  το λόγο του αριθμού των 1-συνόλων, 2-συνόλων, και 3-συνόλων που υπολογίζονται από την ημι-τοπική φάση βελτιστοποίησης προς  $|\mathcal{OPT}|$ , αντίστοιχα. Οι περιορισμένες φάσεις εγγυώνται ότι ο αριθμός των 1-συνόλων στην τελική λύση δεν αυξάνεται, και, επομένως,

$$b_1 \leq a_{i,1}, \text{ για } i = 3, \dots, \ell. \quad (2.7)$$

Σύμφωνα με την απόδειξη του Θεωρήματος 1, έχουμε ότι  $b_1 + b_2 \leq a_{3,1} + a_{3,2} + a_{3,3}$  ενώ είναι σαφές ότι  $b_1 + 2b_2 + 3b_3 = a_{3,1} + 2a_{3,2} + 3a_{3,3}$ . Καταλήγουμε

ότι ο αριθμός  $t_3$  των συνόλων που υπολογίζονται κατά την ημι-τοπική φάση βελτιστοποίησης του αλγορίθμου PRSLI $_{k,\ell}$  είναι

$$\begin{aligned}
t_3 &= (b_1 + b_2 + b_3)|OPT| \\
&\leq \left( \frac{b_1}{3} + \frac{b_1 + b_2}{3} + \frac{b_1 + 2b_2 + 3b_3}{3} \right) |OPT| \\
&\leq \left( \frac{1}{3}b_1 + \frac{2}{3}a_{3,1} + a_{3,2} + \frac{4}{3}a_{3,3} \right) |OPT|. \tag{2.8}
\end{aligned}$$

Με επιχειρήματα παραπλήσια με πριν, έχουμε ότι (2.4) δίνει τον αριθμό  $t_i$  των  $i$ -συνόλων που υπολογίζονται κατά τη φάση ομαδοποίησης ή την περιορισμένη φάση που σχετίζεται με  $i = 4, \dots, k$ . Από (2.4) και (2.8), έχουμε ότι η εγγύηση απόδοσης του αλγορίθμου PRSLI $_{k,\ell}$  είναι

$$\begin{aligned}
\frac{\sum_{i=3}^k t_i}{|OPT|} &\leq \sum_{i=4}^k \frac{1}{i} \left( \sum_{j=1}^i j a_{i,j} - \sum_{j=1}^{i-1} j a_{i-1,j} \right) + \frac{2}{3}a_{3,1} + a_{3,2} + \frac{4}{3}a_{3,3} + \frac{1}{3}b_1 \\
&= \frac{1}{k} \sum_{j=1}^k j a_{k,j} + \sum_{i=4}^{k-1} \frac{1}{i(i+1)} \sum_{j=1}^i j a_{i,j} + \frac{5}{12}a_{3,1} + \frac{1}{2}a_{3,2} \\
&\quad + \frac{7}{12}a_{3,3} + \frac{1}{3}b_1. \tag{2.9}
\end{aligned}$$

Επομένως, ένα άνω φράγμα για το λόγο προσέγγισης του αλγορίθμου PRSLI $_{k,\ell}$  προκύπτει από τη μεγιστοποίηση του δεξιού μέρους της (2.9) υπό τους περιορισμούς (2.1) για  $i = 3, \dots, k$ , (2.3) για  $i = 4, \dots, k$ , και (2.7), με μεταβλητές  $a_{i,j} \geq 0$  για  $i = 3, \dots, k$  και  $j = 1, \dots, i$ , και  $b_1 \geq 0$ . Τυπικά, έχουμε αποδείξει την εξής δήλωση.

**Λήμμα 8.** Ο λόγος προσέγγισης του αλγορίθμου PRSLI $_{k,\ell}$  όταν ένας  $\rho_i$ -προσεγγιστικός αλγόριθμος για  $i$ -ομαδοποίηση συνόλων χρησιμοποιείται κατά τη φάση  $i$  για  $i = \ell + 1, \dots, k$  φράσσεται άνω από τη μέγιστη αντικειμενική τιμή του παρακάτω

γραμμικού προγράμματος:

$$\begin{aligned}
\text{maximize} \quad & \frac{1}{k} \sum_{j=1}^k j a_{k,j} + \sum_{i=4}^{k-1} \frac{1}{i(i+1)} \sum_{j=1}^i j a_{i,j} + \frac{5}{12} a_{3,1} + \frac{1}{2} a_{3,2} + \frac{7}{12} a_{3,3} + \frac{1}{3} b_1 \\
\text{subject to} \quad & \sum_{j=1}^i a_{i,j} \leq 1, i = 3, \dots, k \\
& \sum_{j=1}^{i-1} j a_{i-1,j} - \sum_{j=1}^{i-1} j a_{i,j} - i(1 - \rho_i) a_{i,i} \leq 0, i = \ell + 1, \dots, k \\
& \sum_{j=1}^{i-1} j a_{i-1,j} - \sum_{j=1}^{i-1} j a_{i,j} - (i-1) a_{i,i} \leq 0, i = 4, \dots, \ell \\
& b_1 - a_{i,1} \leq 0, i = 3, \dots, \ell \\
& a_{i,j} \geq 0, i = 3, \dots, k, j = 1, \dots, i \\
& b_1 \geq 0
\end{aligned}$$

Προχωράμε τώρα στην απόδειξη των κεντρικών αποτελεσμάτων μας. Στις δύο επόμενες ενότητες, αποδεικνύουμε ότι οι μέγιστες αντικειμενικές τιμές των γραμμικών προγραμμάτων οι οποίες ‘αποκαλύπτουν’ το λόγο προσέγγισης των αλγορίθμων  $\text{PSL}_{k,4}$  και  $\text{PRSL}_{k,5}$  φράσσονται άνω από τις τιμές που αναφέρονται στα Θεωρήματα 5 και 6. Για να το αποδείξουμε, αρκεί να βρούμε εφικτές λύσεις για τα δυικά γραμμικά προγράμματα που έχουν αυτές τις τιμές σαν αντικειμενικές τιμές τους.



## 2.5 Απόδειξη του Θεωρήματος 5

Το δυικό του γραμμικού προγράμματος του οποίου η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του αλγορίθμου PSLI<sub>k,4</sub> είναι:

$$\begin{aligned}
 & \text{minimize} && \sum_{i=4}^k \beta_i \\
 & \text{subject to} && \beta_4 + \gamma_5 \geq \frac{4}{5} \\
 & && \beta_4 + 2\gamma_5 \geq \frac{3}{5} \\
 & && \beta_4 + 3\gamma_5 \geq \frac{11}{15} \\
 & && \beta_4 + 4\gamma_5 \geq \frac{13}{15} \\
 & && \beta_i + j\gamma_{i+1} - j\gamma_i \geq \frac{j}{i(i+1)}, i = 5, \dots, k-1, j = 1, \dots, i-1 \\
 & && \beta_i + i\gamma_{i+1} - (i-2+\epsilon)\gamma_i \geq \frac{1}{i+1}, i = 5, \dots, k-1 \\
 & && \beta_k - j\gamma_k \geq \frac{j}{k}, j = 1, \dots, k-1 \\
 & && \beta_k - (k-2+\epsilon)\gamma_k \geq 1 \\
 & && \beta_i \geq 0, i = 4, \dots, k \\
 & && \gamma_i \geq 0, i = 5, \dots, k
 \end{aligned}$$

Για να έχουμε μια ιδέα για το πώς είναι μια εφικτή λύση με αντικειμενική τιμή πολύ κοντά στη βέλτιστη για το παραπάνω γραμμικό πρόγραμμα, το λύνουμε χρησιμοποιώντας Matlab για αρκετές τιμές των  $k$  και  $\epsilon = 0$ . Σε όλα τα σχετικά πειράματα, παρατηρούμε ότι ο τελευταίος περιορισμός για τα  $\beta_i$  γίνεται αυστηρός. Οπότε, στην απόδειξη, θέτουμε τις τιμές των  $\beta_i$  κατά τέτοιο τρόπο ώστε οι περιορισμοί αυτοί να είναι αυστηροί ενώ οι τιμές των  $\gamma_i$  είναι τέτοιες ώστε οι υπόλοιποι περιορισμοί του γραμμικού προγράμματος να ικανοποιούνται.

Στην απόδειξη, διακρίνουμε δύο περιπτώσεις ανάλογα με το αν το  $k$  είναι άρτιος ή περιττός.

**Περίπτωση 1:  $k$  άρτιος.** Θέτουμε  $\gamma_k = \frac{1}{k(k-1)}$  και  $\gamma_{k-1} = 0$ . Αν  $k \geq 8$ , θέτουμε  $\gamma_i = \gamma_{i+2} + \frac{2}{i(i+1)(i+2)}$  για  $i = 5, \dots, k-2$ . Θέτουμε επίσης  $\beta_k =$

$1 + (k - 2 + \epsilon)\gamma_k$ ,  $\beta_4 = \frac{13}{15} - 4\gamma_5$  και

$$\beta_i = \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i$$

για  $i = 5, \dots, k-1$ , δηλ., έτσι ώστε ο τελευταίος περιορισμός για το  $\beta_i$  για  $i = 4, \dots, k$  να είναι αυστηρός.

Θα δείξουμε ότι όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος ικανοποιούνται. Προφανώς,  $\gamma_i \geq 0$  για  $i = 5, \dots, k$ . Παρατηρούμε ότι  $\gamma_{k-1} + \gamma_k = \frac{1}{k(k-1)}$ . Αν  $k \geq 8$ , από τον ορισμό του  $\gamma_i$  για  $i = 5, \dots, k-2$ , έχουμε ότι

$$\begin{aligned} \gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} &= \gamma_{i+1} + \gamma_{i+2} + \frac{2}{i(i+1)(i+2)} - \frac{1}{i(i+1)} \\ &= \gamma_{i+1} + \gamma_{i+2} - \frac{1}{(i+1)(i+2)} \end{aligned}$$

και, επομένως,

$$\gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} = \gamma_{k-1} + \gamma_k - \frac{1}{k(k-1)} = 0,$$

δηλ.,  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για  $i = 5, \dots, k-1$ . Ο ορισμός των  $\beta_i$  δίνει

$$\begin{aligned} \beta_i &= \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i \\ &= \frac{i-1}{i(i+1)} - (i-1)\gamma_{i+1} + (i-1)\gamma_i + \epsilon\gamma_i \quad (2.10) \\ &\geq \frac{j}{i(i+1)} - j\gamma_{i+1} + j\gamma_i \\ &\geq 0 \end{aligned}$$

για  $i = 5, \dots, k-1$  και  $j = 1, \dots, i-1$ . Παρατηρούμε ότι η πρώτη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_i$  ικανοποιείται για  $j = 1, \dots, i-1$ . Επομένως, όλοι οι περιορισμοί για το  $\beta_i$  για  $i = 5, \dots, k-1$  ικανοποιούνται.

Οι περιορισμοί για τα  $\beta_k$  επίσης διατηρούνται. Αφού  $\gamma_k = \frac{1}{k(k-1)} \leq \frac{1}{k}$  έχουμε ότι

$$\begin{aligned} \beta_k &= 1 + (k-2+\epsilon)\gamma_k \\ &\geq \frac{k-1}{k} + (k-1)\gamma_k + \epsilon\gamma_k \\ &\geq \frac{j}{k} + j\gamma_k \\ &\geq 0 \end{aligned}$$

για  $j = 1, \dots, k-1$ . Η δεύτερη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_k$  ικανοποιείται για  $j = 1, \dots, k-1$ .

Μένει να αποδείξουμε ότι οι περιορισμοί για το  $\beta_4$  επίσης ικανοποιούνται. Αρκεί να δείξουμε ότι  $\gamma_5 \leq 1/45$ . Αυτό είναι σαφές όταν  $k = 6$ . Αν  $k \geq 8$ , θεωρούμε τις ισότητες  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για περιττό  $i = 5, \dots, k-3$  και  $-\gamma_i - \gamma_{i+1} = -\frac{1}{i(i+1)}$  για άρτιο  $i = 6, \dots, k-2$ . Αθροίζοντας, και αφού  $\gamma_{k-1} = 0$ , έχουμε ότι

$$\begin{aligned}
\gamma_5 &= \sum_{i=3}^{k/2-1} (\gamma_{2i-1} + \gamma_{2i} - \gamma_{2i} - \gamma_{2i+1}) + \gamma_{k-1} \\
&= \sum_{i=3}^{k/2-1} \left( \frac{1}{2i(2i-1)} - \frac{1}{2i(2i+1)} \right) \\
&= \sum_{i=3}^{k/2-1} \left( \frac{1}{2i-1} - \frac{1}{2i} - \frac{1}{2i} + \frac{1}{2i+1} \right) \\
&= \sum_{i=3}^{k/2-1} \left( \frac{1}{2i-1} + \frac{1}{2i+1} \right) - \sum_{i=3}^{k/2-1} \frac{1}{i} \\
&= -\frac{1}{5} + \sum_{i=3}^{k/2-1} \frac{2}{2i-1} + \frac{1}{k-1} - H_{k/2-1} + \frac{3}{2} \\
&= -\frac{1}{5} + 2H_{k-2} - H_{k/2-1} - \frac{8}{3} + \frac{1}{k-1} - H_{k/2-1} + \frac{3}{2} \\
&= 2H_{k-2} - 2H_{k/2-1} + \frac{1}{k-1} - \frac{41}{30} \\
&\leq 2 \ln 2 - \frac{41}{30} \\
&\leq 1/45
\end{aligned}$$

Η πρώτη ανισότητα ισχύει αφού  $2H_{k-2} - 2H_{k/2-1} + \frac{1}{k-1}$  αυξάνεται ως προς  $k$  (αυτό διαπιστώνεται εύκολα εξετάζοντας τον αρχικό ορισμό οφ  $\gamma_5$ ) και αφού  $\lim_{t \rightarrow \infty} H_t / \ln t = 1$ .

Έχουμε αποδείξει ότι όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος ικανοποιούνται, δηλ., η λύση είναι εφικτή. Για να υπολογίσουμε την αντικειμενική τιμή, χρησιμοποιούμε τον ορισμό των  $\beta_i$  και την ισότητα

(2.10). Έχουμε

$$\begin{aligned}
\sum_{i=4}^k \beta_i &= \beta_4 + \beta_5 + \sum_{i=3}^{k/2-1} (\beta_{2i} + \beta_{2i+1}) + \beta_k \\
&= \frac{13}{15} - 4\gamma_5 + \frac{2}{15} + 4\gamma_5 - 4\gamma_6 + \sum_{i=3}^{k/2-1} \left( \frac{1}{2i+1} - 2i\gamma_{2i+1} + (2i-2)\gamma_{2i} \right. \\
&\quad \left. + \frac{2i}{(2i+1)(2i+2)} - 2i\gamma_{2i+2} + 2i\gamma_{2i+1} \right) + 1 + (k-2)\gamma_k + \epsilon \sum_{i=5}^k \gamma_i \\
&= 2 + \sum_{i=3}^{k/2-1} \frac{1}{i+1} - 4\gamma_6 + \sum_{i=3}^{k/2-1} ((2i-2)\gamma_{2i} - 2i\gamma_{2i+2}) + (k-2)\gamma_k \\
&\quad + \epsilon \sum_{i=5}^k \gamma_i \\
&\leq H_{k/2} + 1/6 + \epsilon
\end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει αφού

$$\sum_{i=5}^k \gamma_i \leq \sum_{i=5}^k \frac{1}{i(i-1)} = \sum_{i=5}^k \left( \frac{1}{i-1} - \frac{1}{i} \right) = 1/4 - 1/k.$$

**Περίπτωση 2:  $k$  περιττός.** Θέτουμε

$$\gamma_5 = \min \left\{ \frac{1}{45}, 2H_k - 2H_{\frac{k-1}{2}} - \frac{41}{30} + \frac{1}{k-1} - \frac{2}{k} \right\},$$

δηλ.,  $\gamma_5 = 1/45$  για  $k \in \{5, 7, 9, 11, 13\}$  και  $\gamma_5 = 2H_k - 2H_{\frac{k-1}{2}} - \frac{41}{30} + \frac{1}{k-1} - \frac{2}{k}$ , διαφορετικά. Αν  $k \geq 7$ , θέτουμε  $\gamma_6 = \frac{1}{30} - \gamma_5$  και  $\gamma_i = \gamma_{i-2} - \frac{2}{i(i-1)(i-2)}$  για  $i = 7, \dots, k$ . Θέτουμε επίσης  $\beta_k = 1 + (k-2 + \epsilon)\gamma_k$ ,  $\beta_4 = \frac{13}{15} - 4\gamma_5$  και, αν  $k \geq 7$ ,

$$\beta_i = \frac{1}{i+1} - i\gamma_{i+1} + (i-2 + \epsilon)\gamma_i$$

για  $i = 5, \dots, k-1$ , δηλ., έτσι ώστε ο τελευταίος περιορισμός για το  $\beta_i$  για  $i = 4, \dots, k$  να γίνεται αυστηρός.

Θα δείξουμε ότι ικανοποιούνται όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος. Αν  $k \geq 7$ , παρατηρούμε ότι  $\gamma_5 + \gamma_6 = \frac{1}{30}$ . Από τον ορισμό του  $\gamma_i$  για  $i = 7, \dots, k-1$ , έχουμε

$$\begin{aligned} \gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} &= \gamma_{i-1} + \gamma_i - \frac{2}{(i-1)i(i+1)} - \frac{1}{i(i+1)} \\ &= \gamma_{i-1} + \gamma_i - \frac{1}{i(i-1)} \end{aligned}$$

και, επομένως,

$$\gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} = \gamma_5 + \gamma_6 - \frac{1}{30} = 0,$$

δηλ.,  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για  $i = 5, \dots, k-1$ . Ο ορισμός των  $\beta_i$  δίνει

$$\begin{aligned} \beta_i &= \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i \\ &= \frac{i-1}{i(i+1)} - (i-1)\gamma_{i+1} + (i-1)\gamma_i + \epsilon\gamma_i \quad (2.11) \\ &\geq \frac{j}{i(i+1)} - j\gamma_{i+1} + j\gamma_i \\ &\geq 0 \end{aligned}$$

για  $i = 5, \dots, k-1$  και  $j = 1, \dots, i-1$ . Παρατηρούμε ότι η πρώτη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_i$  ικανοποιείται για  $j = 1, \dots, i-1$ . Οι περιορισμοί για τα  $\beta_4$  ικανοποιούνται αφού  $\gamma_5 \leq 1/45$ . Επομένως, όλοι οι περιορισμοί για τα  $\beta_i$  για  $i = 4, \dots, k-1$  ικανοποιούνται.

Δείχνουμε τώρα ότι  $\gamma_i \geq 0$ , για  $i = 5, \dots, k$  και ότι οι περιορισμοί για το  $\beta_k$  ικανοποιούνται. Αν  $k = 5$ , είναι σαφές. Αν  $k \geq 7$ , από τον ορισμό του  $\gamma_i$ , για να δείξουμε ότι  $\gamma_i \geq 0$  αρκεί να δείξουμε ότι  $\gamma_{k-1} \geq 0$  και  $\gamma_k \geq 0$ . Θεωρούμε τις ισότητες  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για περικό  $i = 5, \dots, k-2$  και

$-\gamma_i - \gamma_{i+1} = -\frac{1}{i(i+1)}$  για άρτιο  $i = 6, \dots, k-3$ . Αθροίζοντας, έχουμε ότι

$$\begin{aligned}
\gamma_{k-1} &= -\gamma_5 + \sum_{i=3}^{\frac{k-3}{2}} (\gamma_{2i-1} + \gamma_{2i} - \gamma_{2i} - \gamma_{2i+1}) + \gamma_{k-2} + \gamma_{k-1} \\
&= -\gamma_5 + \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i(2i-1)} - \frac{1}{2i(2i+1)} \right) + \frac{1}{(k-1)(k-2)} \\
&= -\gamma_5 + \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i-1} - \frac{1}{2i} - \frac{1}{2i} + \frac{1}{2i+1} \right) + \frac{1}{k-2} - \frac{1}{k-1} \\
&= -\gamma_5 + \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i-1} + \frac{1}{2i+1} \right) - \sum_{i=3}^{\frac{k-1}{2}} \frac{1}{i} + \frac{1}{k-2} + \frac{1}{k-1} \\
&= -\gamma_5 - \frac{1}{5} + \sum_{i=3}^{\frac{k-1}{2}} \frac{2}{2i-1} - \sum_{i=3}^{\frac{k-1}{2}} \frac{1}{i} + \frac{1}{k-1} \\
&= -\gamma_5 - \frac{1}{5} + 2H_{k-2} - H_{\frac{k-3}{2}} - \frac{8}{3} - H_{\frac{k-1}{2}} + \frac{3}{2} + \frac{1}{k-1} \\
&= -\gamma_5 + 2H_k - 2H_{\frac{k-1}{2}} - \frac{41}{30} + \frac{1}{k-1} - \frac{2}{k}.
\end{aligned}$$

Από τον ορισμό του  $\gamma_5$ , έχουμε ότι  $\gamma_{k-1} \geq 0$ . Συγκεκριμένα,  $\gamma_{k-1} = 0$  για  $k \geq 15$  ενώ, παρατηρώντας, έχουμε ότι  $\gamma_{k-1} \leq \frac{1}{k(k-1)}$  για  $k \in \{5, 7, 9, 11, 13\}$  και, επομένως,  $\gamma_k = \frac{1}{k(k-1)} - \gamma_{k-1} \in \left[0, \frac{1}{k(k-1)}\right]$ . Επομένως, αφού  $\gamma_k \leq \frac{1}{k}$ , έχουμε ότι

$$\begin{aligned}
\beta_k &= 1 + (k-2 + \epsilon)\gamma_k \\
&\geq \frac{k-1}{k} + (k-1)\gamma_k + \epsilon\gamma_k \\
&\geq \frac{j}{k} + j\gamma_k \\
&\geq 0
\end{aligned}$$

για  $j = 1, \dots, k-1$ . Η δεύτερη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_k$  ικανοποιείται για  $j = 1, \dots, k-1$ .

Έχουμε αποδείξει ότι όλοι οι περιορισμοί του δικού γραμμικού προγράμματος ικανοποιούνται, δηλ., η λύση είναι εφικτή. Για να υπολογίσουμε

την αντικειμενική τιμή, χρησιμοποιούμε τον ορισμό των  $\beta_i$  και την ισότητα (2.11). Έχουμε

$$\begin{aligned}
\sum_{i=4}^k \beta_i &= \beta_4 + \sum_{i=3}^{\frac{k-1}{2}} (\beta_{2i-1} + \beta_{2i}) + \beta_k \\
&= \frac{13}{15} - 4\gamma_5 + \sum_{i=3}^{\frac{k-1}{2}} \left( \frac{1}{2i} - (2i-1)\gamma_{2i} + (2i-3)\gamma_{2i-1} + \frac{2i-1}{2i(2i+1)} \right. \\
&\quad \left. - (2i-1)\gamma_{2i+1} + (2i-1)\gamma_{2i} \right) + 1 + (k-2)\gamma_k + \epsilon \sum_{i=5}^k \gamma_i \\
&= \frac{28}{15} + \sum_{i=3}^{\frac{k-1}{2}} \frac{2}{2i+1} - 4\gamma_5 + \sum_{i=3}^{\frac{k-1}{2}} ((2i-3)\gamma_{2i-1} - (2i-1)\gamma_{2i+1}) \\
&\quad + (k-2)\gamma_k + \epsilon \sum_{i=5}^k \gamma_i \\
&\leq -\frac{6}{5} - \gamma_5 + 2H_k - H_{\frac{k-1}{2}} + \epsilon
\end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει αφού

$$\sum_{i=5}^k \gamma_i \leq \sum_{i=5}^k \frac{1}{i(i-1)} = \sum_{i=5}^k \left( \frac{1}{i-1} - \frac{1}{i} \right) = 1/4 - 1/k.$$

Τώρα, αφού  $\gamma_5 = 1/45$  όταν  $k \in \{5, 7, 9, 11, 13\}$ , έχουμε ότι

$$\sum_{i=4}^k \beta_i \leq 2H_k - H_{\frac{k-1}{2}} - \frac{11}{9} + \epsilon.$$

Αν  $k \geq 15$ , έχουμε  $\gamma_5 = 2H_k - 2H_{\frac{k-1}{2}} - \frac{41}{30} + \frac{1}{k-1} - \frac{2}{k}$  και, επομένως,

$$\sum_{i=4}^k \beta_i \leq H_{\frac{k-1}{2}} + \frac{1}{6} + \frac{2}{k} - \frac{1}{k-1} + \epsilon.$$

Λόγω της δυσκολίας,  $\sum_{i=4}^k \beta_i$  είναι ένα άνω φράγμα για τη μέγιστη αντικειμενική τιμή του γραμμικού προγράμματος η οποία 'αποκαλύπτει' το λόγο προσέγγισης του αλγορίθμου. Το θεώρημα προκύπτει από το Λήμμα 7.  $\square$

## 2.6 Απόδειξη του Θεωρήματος 6

Το δυκό του γραμμικού προγράμματος του οποίου η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του αλγορίθμου PRSLI<sub>k,5</sub> ισ:

$$\begin{aligned}
 & \text{minimize} && \sum_{i=3}^k \beta_i \\
 & \text{subject to} && \delta_3 + \delta_4 + \delta_5 \geq \frac{1}{3} \\
 & && \beta_3 + \gamma_4 - \delta_3 \geq \frac{5}{12} \\
 & && \beta_3 + 2\gamma_4 \geq \frac{1}{2} \\
 & && \beta_3 + 3\gamma_4 \geq \frac{7}{12} \\
 & && \beta_i + \gamma_{i+1} - \gamma_i - \delta_i \geq \frac{1}{i(i+1)}, i = 4, 5 \\
 & && \beta_i + j\gamma_{i+1} - j\gamma_i \geq \frac{j}{i(i+1)}, i = 4, 5, j = 2, \dots, i-1 \\
 & && \beta_i + i\gamma_{i+1} - (i-1)\gamma_i \geq \frac{1}{i+1}, i = 4, 5 \\
 & && \beta_i + j\gamma_{i+1} - j\gamma_i \geq \frac{j}{i(i+1)}, i = 6, \dots, k-1, j = 1, \dots, i-1 \\
 & && \beta_i + i\gamma_{i+1} - (i-2+\epsilon)\gamma_i \geq \frac{1}{i+1}, i = 6, \dots, k-1 \\
 & && \beta_k - j\gamma_k \geq \frac{j}{k}, j = 1, \dots, k-1 \\
 & && \beta_k - (k-2+\epsilon)\gamma_k \geq 1 \\
 & && \beta_i \geq 0, i = 3, \dots, k \\
 & && \gamma_i \geq 0, i = 4, \dots, k \\
 & && \delta_i \geq 0, i = 3, \dots, 5
 \end{aligned}$$

Όπως είναι αναμενόμενο, το δυκό του γραμμικού προγράμματος του οποίου η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του αλγορίθμου PRSLI<sub>k,5</sub> είναι πιο πολύπλοκο από αυτό για τον αλγόριθμο PSII<sub>k,4</sub>. Ενώ οι περιορισμοί για το  $\beta_i$  για  $i \geq 6$  είναι ίδιοι με πριν, οι υπόλοιποι περιορισμοί είναι διαφορετικοί. Επιπλέον, η ύπαρξη των μεταβλητών  $\delta_3, \delta_4$ , και



$\delta_5$  κάνει ακόμα πιο πολύπλοκο το γραμμικό πρόγραμμα. Στην απόδειξη, έχουμε χρησιμοποιήσει σαν οδηγό τη λύση στο γραμμικό πρόγραμμα για αρκετές τιμές των  $k$  και  $\epsilon = 0$  χρησιμοποιώντας Matlab. Σε όλα τα πειράματα, παρατηρήσαμε ότι ο τελευταίος περιορισμός για το  $\beta_i$  για  $i \geq 6$  να γίνεται αυστηρός. Επιπλέον, οι περισσότεροι από τους περιορισμούς για τα  $\beta_3, \beta_4$ , και  $\beta_5$  είναι επίσης αυστηροί. Έτσι, στην απόδειξη, θέτουμε τις τιμές των  $\beta_i$  έτσι ώστε οι περιορισμοί αυτοί να είναι αυστηροί ενώ δίνουμε κατάλληλες τιμές στα  $\gamma_i$  και  $\delta_3, \delta_4$ , και  $\delta_5$  έτσι ώστε οι υπόλοιποι περιορισμοί του γραμμικού προγράμματος να ικανοποιούνται.

Στην απόδειξη, διακρίνουμε δύο περιπτώσεις ανάλογα με το αν ο  $k$  είναι περιττός ή άρτιος.

**Περίπτωση 1:  $k$  περιττός.** Θέτουμε  $\gamma_4 = 1/12$ ,  $\gamma_5 = 1/20$ ,  $\gamma_k = \frac{1}{k(k-1)}$  και  $\gamma_{k-1} = 0$ . Αν  $k \geq 9$ , θέτουμε  $\gamma_i = \gamma_{i+2} + \frac{2}{i(i+1)(i+2)}$  για  $i = 6, \dots, k-2$ . Θέτουμε επίσης  $\delta_3 = 0$ ,  $\delta_4 = 1/6$ ,  $\delta_5 = 17/60 - 4\gamma_6$ ,  $\beta_3 = 1/3$ ,  $\beta_4 = 1/4$ ,  $\beta_5 = 11/30 - 5\gamma_6$ ,  $\beta_k = 1 + (k-2+\epsilon)\gamma_k$ , και

$$\beta_i = \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i$$

για  $i = 6, \dots, k-1$ .

Θα δείξουμε ότι ικανοποιούνται όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος. Προφανώς,  $\gamma_i \geq 0$  για  $i = 5, \dots, k$ . Παρατηρούμε ότι  $\gamma_{k-1} + \gamma_k = \frac{1}{k(k-1)}$ . Αν  $k \geq 9$ , από τον ορισμό του  $\gamma_i$  για  $i = 6, \dots, k-2$ , έχουμε

$$\begin{aligned} \gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} &= \gamma_{i+1} + \gamma_{i+2} + \frac{2}{i(i+1)(i+2)} - \frac{1}{i(i+1)} \\ &= \gamma_{i+1} + \gamma_{i+2} - \frac{1}{(i+1)(i+2)} \end{aligned}$$

και, επομένως,

$$\gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} = \gamma_{k-1} + \gamma_{k-2} - \frac{1}{(k-1)(k-2)} = 0,$$

δηλ.,  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για  $i = 6, \dots, k-1$ . Ο ορισμός των  $\beta_i$  δίνει

$$\begin{aligned}
\beta_i &= \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i \\
&= \frac{i-1}{i(i+1)} - (i-1)\gamma_{i+1} + (i-1)\gamma_i + \epsilon\gamma_i & (2.12) \\
&\geq \frac{j}{i(i+1)} - j\gamma_{i+1} + j\gamma_i \\
&\geq 0
\end{aligned}$$

για  $i = 6, \dots, k-1$  και  $j = 1, \dots, i-1$ . Παρατηρούμε ότι η πρώτη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_i$  ικανοποιείται για  $j = 1, \dots, i-1$ . Επομένως, όλοι οι περιορισμοί για τα  $\beta_i$  για  $i = 6, \dots, k-1$  ικανοποιούνται.

Οι περιορισμοί για τα  $\beta_k$  επίσης ικανοποιούνται. Αφού  $\gamma_k = \frac{1}{k(k-1)} \leq \frac{1}{k}$  έχουμε ότι

$$\begin{aligned}
\beta_k &= 1 + (k-2+\epsilon)\gamma_k \\
&\geq \frac{k-1}{k} + (k-1)\gamma_k + \epsilon\gamma_k \\
&\geq \frac{j}{k} + j\gamma_k \\
&\geq 0
\end{aligned}$$

για  $j = 1, \dots, k-1$ . Οι περιορισμοί για τα  $\beta_3$  και  $\beta_4$  επαληθεύονται εύκολα μέσω απλών υπολογισμών. Η δεύτερη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_k$  ικανοποιείται για  $j = 1, \dots, k-1$ .

Για να δείξουμε ότι και οι υπόλοιποι περιορισμοί ικανοποιούνται, θα δείξουμε πρώτα ότι  $\gamma_6 \leq 7/240$ . Αυτό είναι σαφές όταν  $k = 7$ . Αν  $k \geq 9$ , θεωρούμε τις ισότητες  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για άρτιο  $i = 6, \dots, k-3$  και  $-\gamma_i - \gamma_{i+1} = -\frac{1}{i(i+1)}$  για περιττό  $i = 7, \dots, k-2$ . Αθροίζοντας, και αφού

$\gamma_{k-1} = 0$ , έχουμε ότι

$$\begin{aligned}
\gamma_6 &= \sum_{i=3}^{\frac{k-3}{2}} (\gamma_{2i} + \gamma_{2i+1} - \gamma_{2i+1} - \gamma_{2i+2}) + \gamma_{k-1} \\
&= \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i(2i+1)} - \frac{1}{(2i+1)(2i+2)} \right) \\
&= \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i} - \frac{1}{2i+1} - \frac{1}{2i+1} + \frac{1}{2i+2} \right) \\
&= \sum_{i=3}^{\frac{k-3}{2}} \left( \frac{1}{2i} + \frac{1}{2i+2} \right) - \sum_{i=3}^{\frac{k-3}{2}} \frac{2}{2i+1} \\
&= -\frac{1}{6} + H_{\frac{k-3}{2}} - \frac{3}{2} + \frac{1}{k-1} - 2H_{k-2} + H_{\frac{k-3}{2}} + 2 + \frac{2}{3} + \frac{2}{5} \\
&= 2H_{\frac{k-3}{2}} - 2H_{k-2} + \frac{1}{k-1} + \frac{7}{5} \\
&\leq -2\ln 2 + \frac{7}{5} \\
&\leq 7/240
\end{aligned}$$

Η πρώτη ανισότητα ισχύει αφού  $2H_{\frac{k-3}{2}} - 2H_{k-2} + \frac{1}{k-1}$  αυξάνεται ως προς  $k$  (αυτό διαπιστώνεται εύκολα εξετάζοντας τον αρχικό ορισμό του  $\gamma_6$ ) και αφού  $\lim_{t \rightarrow \infty} H_t / \ln t = 1$ .

Αφού  $\gamma_6 \leq 7/240$ , έχουμε  $\delta_3 + \delta_4 + \delta_5 = 1/6 + 17/60 - 4\gamma_6 \geq 1/3$ . Όσον αφορά στους περιορισμούς για το  $\beta_5$ , έχουμε  $\beta_5 + \gamma_6 - \gamma_5 - \delta_5 = 11/30 - 5\gamma_6 + \gamma_6 - 1/20 - 17/60 + 4\gamma_6 = 1/30$ ,  $\beta_5 + 5\gamma_6 - 4\gamma_5 = \frac{11}{30} - 5\gamma_6 + \gamma_6 - \frac{1}{5} = \frac{1}{6}$  και, για  $j = 2, 3, 4$ ,

$$\begin{aligned}
\beta_5 &= \frac{11}{30} - 5\gamma_6 \\
&\geq \frac{11}{30} - (5-j)\frac{7}{240} - j\gamma_6 \\
&= j \left( \frac{1}{30} - \gamma_6 + \frac{1}{20} \right) - \frac{13j}{240} + \frac{53}{240} \\
&\geq j \left( \frac{1}{5 \cdot 6} - \gamma_6 + \gamma_5 \right).
\end{aligned}$$

Τέλος, το γεγονός ότι  $\gamma_6 \leq 7/240$  επίσης εγγυάται ότι  $\delta_5$  και  $\beta_5$  είναι μη αρνητικοί.

Έχουμε αποδείξει ότι όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος ικανοποιούνται, δηλ., η λύση είναι εφικτή. Για να υπολογίσουμε την αντικειμενική τιμή, χρησιμοποιούμε τον ορισμό των  $\beta_i$  και την ισότητα (2.12). Έχουμε

$$\begin{aligned}
\sum_{i=3}^k \beta_i &= \beta_3 + \beta_4 + \beta_5 + \beta_6 + \sum_{i=4}^{\frac{k-1}{2}} (\beta_{2i-1} + \beta_{2i}) + \beta_k \\
&= \frac{1}{3} + \frac{1}{4} + \frac{11}{30} - 5\gamma_6 + \frac{5}{42} - 5\gamma_7 + 5\gamma_6 + \sum_{i=4}^{\frac{k-1}{2}} \left( \frac{1}{2i} - (2i-1)\gamma_{2i} \right. \\
&\quad \left. + (2i-3)\gamma_{2i-1} + \frac{2i-1}{2i(2i+1)} - (2i-1)\gamma_{2i+1} + (2i-1)\gamma_{2i} \right) \\
&\quad + 1 + (k-2)\gamma_k + \epsilon \sum_{i=6}^k \gamma_i \\
&= \frac{869}{420} + \sum_{i=4}^{\frac{k-1}{2}} \frac{2}{2i+1} - 5\gamma_7 + \sum_{i=4}^{\frac{k-1}{2}} ((2i-3)\gamma_{2i-1} - (2i-1)\gamma_{2i+1}) \\
&\quad + (k-2)\gamma_k + \epsilon \sum_{i=6}^k \gamma_i \\
&\leq 2H_k - H_{\frac{k-1}{2}} - 77/60 + \epsilon
\end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει αφού

$$\sum_{i=6}^k \gamma_i \leq \sum_{i=6}^k \frac{1}{i(i-1)} = \sum_{i=6}^k \left( \frac{1}{i-1} - \frac{1}{i} \right) = 1/5 - 1/k.$$

**Περίπτωση 2:  $k$  άρτιο.** Θέτουμε  $\gamma_4 = 1/12$ ,  $\gamma_5 = 1/20$ , και

$$\gamma_6 = \min \left\{ \frac{7}{240}, \frac{7}{5} - 2H_k + 2H_{k/2} + \frac{1}{k-1} - \frac{2}{k} \right\},$$

δηλ.,  $\gamma_6 = 7/240$  για  $k = 6$  και  $\gamma_6 = \frac{7}{5} - 2H_k + 2H_{k/2} + \frac{1}{k-1} - \frac{2}{k}$ , διαφορετικά. Αν  $k \geq 8$ , θέτουμε  $\gamma_7 = \frac{1}{42} - \gamma_6$  και  $\gamma_i = \gamma_{i-2} + \frac{2}{i(i-1)(i-2)}$  για  $i = 8, \dots, k$ .

Θέτουμε επίσης  $\delta_3 = 0$ ,  $\delta_4 = 1/6$ ,  $\delta_5 = 17/60 - 4\gamma_6$ ,  $\beta_3 = 1/3$ ,  $\beta_4 = 1/4$ ,  $\beta_5 = 11/30 - 5\gamma_6$ ,  $\beta_k = 1 + (k - 2 + \epsilon)\gamma_k$ , και, αν  $k \geq 8$ ,

$$\beta_i = \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i$$

για  $i = 6, \dots, k-1$ .

Θα δείξουμε ότι ικανοποιούνται όλοι οι περιορισμοί του δυϊκού γραμμικού προγράμματος. Αν  $k \geq 8$ , παρατηρούμε ότι  $\gamma_6 + \gamma_7 = \frac{1}{42}$ . Από τον ορισμό του  $\gamma_i$  για  $i = 8, \dots, k$ , έχουμε

$$\begin{aligned} \gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} &= \gamma_{i-1} + \gamma_i - \frac{2}{(i-1)i(i+1)} - \frac{1}{i(i+1)} \\ &= \gamma_{i-1} + \gamma_i - \frac{1}{i(i-1)} \end{aligned}$$

και, επομένως,

$$\gamma_i + \gamma_{i+1} - \frac{1}{i(i+1)} = \gamma_6 + \gamma_7 - \frac{1}{42} = 0,$$

δηλ.,  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για  $i = 6, \dots, k-1$ . Ο ορισμός των  $\beta_i$  δίνει

$$\begin{aligned} \beta_i &= \frac{1}{i+1} - i\gamma_{i+1} + (i-2+\epsilon)\gamma_i \\ &= \frac{i-1}{i(i+1)} - (i-1)\gamma_{i+1} + (i-1)\gamma_i + \epsilon\gamma_i \quad (2.13) \\ &\geq \frac{j}{i(i+1)} - j\gamma_{i+1} + j\gamma_i \\ &\geq 0 \end{aligned}$$

για  $i = 6, \dots, k-1$  και  $j = 1, \dots, i-1$ . Παρατηρούμε ότι η πρώτη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_i$  ικανοποιείται για  $j = 1, \dots, i-1$ . Οι περιορισμοί για τα  $\beta_3$  και  $\beta_4$  εύκολα επαληθεύονται έσω απλών υπολογισμών. Για τους περιορισμούς για το  $\beta_5$ , έχουμε  $\beta_5 + \gamma_6 - \gamma_5 - \delta_5 = 11/30 - 5\gamma_6 + \gamma_6 - 1/20 - 17/60 + 4\gamma_6 = 1/30$ ,  $\beta_5 + 5\gamma_6 - 4\gamma_5 = \frac{11}{30} - 5\gamma_6 + \gamma_6 - \frac{1}{5} = \frac{1}{6}$ , και,

για  $j = 2, 3, 4$ ,

$$\begin{aligned}
\beta_5 &= \frac{11}{30} - 5\gamma_6 \\
&\geq \frac{11}{30} - (5-j)\frac{7}{240} - j\gamma_6 \\
&= j\left(\frac{1}{30} - \gamma_6 + \frac{1}{20}\right) - \frac{13j}{240} + \frac{53}{240} \\
&\geq j\left(\frac{1}{5 \cdot 6} - \gamma_6 + \gamma_5\right).
\end{aligned}$$

Αφού  $\gamma_6 \leq 7/240$ , ισχύει ότι  $\delta_3 + \delta_4 + \delta_5 = 1/6 + 17/60 - 4\gamma_6 \geq 1/3$ . Επίσης, το γεγονός ότι  $\gamma_6 \leq 7/240$  εγγυάται ότι  $\delta_5$  και  $\beta_5$  είναι μη αρνητικοί. Επομένως, όλοι οι περιορισμοί για τα  $\delta_3, \delta_4, \delta_5$  και  $\beta_i$  για  $i = 3, \dots, k-1$  ικανοποιούνται.

Δείχνουμε τώρα ότι  $\gamma_i \geq 0$ , για  $i = 6, \dots, k$  και ότι οι περιορισμοί για το  $\beta_k$  ικανοποιούνται. Αν  $k = 6$ , είναι σαφές. Αν  $k \geq 8$ , από τον ορισμό του  $\gamma_i$ , για να δείξουμε ότι  $\gamma_i \geq 0$  αρκεί να δείξουμε ότι  $\gamma_{k-1} \geq 0$  και  $\gamma_k \geq 0$ . Θεωρούμε τις ισότητες  $\gamma_i + \gamma_{i+1} = \frac{1}{i(i+1)}$  για άρτιο  $i = 6, \dots, k-2$  και  $-\gamma_i - \gamma_{i+1} = -\frac{1}{i(i+1)}$  για περιττό  $i = 7, \dots, k-3$ . Αθροίζοντας, έχουμε ότι

$$\begin{aligned}
\gamma_{k-1} &= -\gamma_6 + \sum_{i=3}^{k/2-2} (\gamma_{2i} + \gamma_{2i+1} - \gamma_{2i+1} - \gamma_{2i+2}) + \gamma_{k-2} + \gamma_{k-1} \\
&= -\gamma_6 + \sum_{i=3}^{k/2-2} \left( \frac{1}{2i(2i+1)} - \frac{1}{(2i+1)(2i+2)} \right) + \frac{1}{(k-1)(k-2)} \\
&= -\gamma_6 + \sum_{i=3}^{k/2-2} \left( \frac{1}{2i} - \frac{1}{2i+1} - \frac{1}{2i+1} + \frac{1}{2i+2} \right) + \frac{1}{k-2} - \frac{1}{k-1} \\
&= -\gamma_6 + \sum_{i=3}^{k/2-2} \left( \frac{1}{2i} + \frac{1}{2i+2} \right) - \sum_{i=3}^{k/2-2} \frac{2}{2i+1} + \frac{1}{k-2} - \frac{1}{k-1} \\
&= -\gamma_6 - \frac{1}{6} + \sum_{i=3}^{k/2-1} \frac{1}{i} - \sum_{i=3}^{k/2-1} \frac{2}{2i+1} + \frac{1}{k-1} \\
&= -\gamma_6 - \frac{1}{6} - \frac{3}{2} + H_{k/2} - \frac{2}{k} + \frac{46}{15} - 2H_{k-1} + H_{k/2-1} + \frac{1}{k-1} \\
&= -\gamma_6 + \frac{7}{5} + 2H_{k/2} - 2H_k + \frac{1}{k-1} - \frac{2}{k}.
\end{aligned}$$

Από τον ορισμό του  $\gamma_6$ , έχουμε ότι  $\gamma_{k-1} = 0$  όταν  $k \geq 8$  και, επομένως,  $\gamma_k = \frac{1}{k(k-1)} - \gamma_{k-1} = \frac{1}{k(k-1)}$ . Επομένως, αφού  $\gamma_k \leq \frac{1}{k}$ , έχουμε ότι

$$\begin{aligned}\beta_k &= 1 + (k - 2 + \epsilon)\gamma_k \\ &\geq \frac{k-1}{k} + (k-1)\gamma_k + \epsilon\gamma_k \\ &\geq \frac{j}{k} + j\gamma_k \\ &\geq 0\end{aligned}$$

για  $j = 1, \dots, k-1$ . Η δεύτερη ανισότητα εγγυάται ότι ο πρώτος περιορισμός για  $b_k$  ικανοποιείται για  $j = 1, \dots, k-1$ .

Έχουμε αποδείξει ότι όλοι οι περιορισμοί του δυικού γραμμικού προγράμματος ικανοποιούνται, δηλ., η λύση είναι εφικτή. Για να υπολογίσουμε την αντικειμενική τιμή, χρησιμοποιούμε τον ορισμό των  $\beta_i$  και την ισότητα (2.13). Έχουμε

$$\begin{aligned}\sum_{i=3}^k \beta_i &= \beta_3 + \beta_4 + \beta_5 + \sum_{i=3}^{k/2-1} (\beta_{2i} + \beta_{2i+1}) + \beta_k \\ &= \frac{1}{3} + \frac{1}{4} + \frac{11}{30} - 5\gamma_6 + \sum_{i=3}^{k/2-1} \left( \frac{1}{2i+1} - 2i\gamma_{2i+1} + (2i-2)\gamma_{2i} \right. \\ &\quad \left. + \frac{2i}{(2i+1)(2i+2)} - 2i\gamma_{2i+2} + 2i\gamma_{2i+1} \right) + 1 + (k-2)\gamma_k + \epsilon \sum_{i=6}^k \gamma_i \\ &= \frac{39}{20} + \sum_{i=3}^{k/2-1} \frac{1}{i+1} - 5\gamma_6 + \sum_{i=3}^{k/2-1} ((2i-2)\gamma_{2i} - 2i\gamma_{2i+2}) + (k-2)\gamma_k \\ &\quad + \epsilon \sum_{i=6}^k \gamma_i \\ &\leq \frac{7}{60} - \gamma_6 + H_{k/2} + \epsilon\end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει αφού

$$\sum_{i=6}^k \gamma_i \leq \sum_{i=6}^k \frac{1}{i(i-1)} = \sum_{i=6}^k \left( \frac{1}{i-1} - \frac{1}{i} \right) = 1/5 - 1/k.$$

Τώρα, αφού  $\gamma_6 = 7/240$  όταν  $k = 6$ , έχουμε ότι

$$\sum_{i=3}^6 \beta_i \leq \frac{461}{240} + \epsilon.$$

Αν  $k \geq 8$ , έχουμε  $\gamma_6 = \frac{7}{5} - 2H_k + 2H_{k/2} + \frac{1}{k-1} - \frac{2}{k}$  και, επομένως,

$$\sum_{i=3}^k \beta_i \leq 2H_k - H_{k/2} - \frac{77}{60} + \frac{2}{k} - \frac{1}{k-1} + \epsilon.$$

Λόγω της δυσκολίας,  $\sum_{i=3}^k \beta_i$  είναι ένα άνω φράγμα για τη μέγιστη αντικειμενική τιμή του του γραμμικού προγράμματος η οποία ‘αποκαλύπτει’ το λόγο προσέγγισης του αλγορίθμου. Το θεώρημα προκύπτει από το Λήμμα 8.  $\square$

## 2.7 Επεκτάσεις

Έχουμε επαληθεύσει πειραματικά με χρήση Matlab ότι τα άνω φράγματα που παρουσιάζουμε είναι αυστηρά με την έννοια ότι αποτελούν τις μέγιστες αντικειμενικές τιμές γραμμικών προγραμμάτων οι οποίες ‘αποκαλύπτουν’ το λόγο προσέγγισης των υπό εξέταση αλγορίθμων (αγνοώντας τον όρο  $\epsilon$  στο φράγμα προσέγγισης). Η τεχνική μας για την ανάλυση μπορεί επίσης να χρησιμοποιηθεί ώστε να δώσει απλές αποδείξεις για τα αποτελέσματα στην εργασία [34] (δηλ., τα Θεωρήματα 2 και 3). Οι αρκετές περιπτώσεις που μελετώνται στις αποδείξεις της εργασίας [34] περιλαμβάνονται στην ουσία σαν περιορισμοί στα γραμμικά προγράμματα των οποίων οι αντικειμενικές τιμές ‘αποκαλύπτουν’ το λόγο προσέγγισης των αλγορίθμων τα οποία είναι πολύ πιο απλά από τα αντίστοιχα για τους αλγόριθμους  $\text{PSLI}_{k,\ell}$  και  $\text{PRSLI}_{k,\ell}$ . Επιπλέον, σημειώνουμε ότι δεν έχουμε συνδυάσει τις τεχνικές μας με τον πρόσφατο αλγόριθμο του Levin [59] που χειρίζεται σύνολα μεγέθους 4 χρησιμοποιώντας μια περιορισμένη φάση τοπικής αναζήτησης. Εικάζουμε ότι είναι δυνατές και περαιτέρω βελτιώσεις.



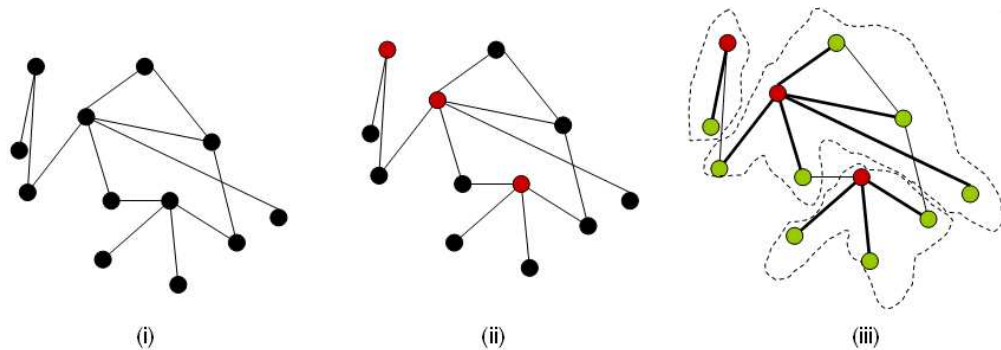
## Κεφάλαιο 3

# Εύρεση γεννητικού δάσους αστέρων και εξοικονόμηση χρωμάτων

Σε αυτό το κεφάλαιο, παρουσιάζουμε ένα απλό αλγόριθμο για το πρόβλημα εύρεσης μέγιστου γεννητικού δάσους αστέρων. Λαμβάνουμε υπόψη το γεγονός ότι το πρόβλημα αποτελεί ειδική περίπτωση του συμπληρωματικού προβλήματος κάλυψης με σύνολα και προσαρμόζουμε έναν αλγόριθμο των Duh και Fűrer για την επίλυσή του. Αποδεικνύουμε ότι ο αλγόριθμος αυτός υπολογίζει  $193/240 \approx 0.804$ -προσεγγιστικά γεννητικά δάση αστέρων. Το αποτέλεσμα αυτό βελτιώνει ένα προηγούμενο κάτω φράγμα με τιμή 0.71 των Chen και άλλων. Αν και ο αλγόριθμος είναι καθαρά συνδυαστικός, η ανάλυσή μας ορίζει ένα γραμμικό πρόγραμμα που χρησιμοποιεί μια παράμετρο  $f$  το οποίο είναι επιλύσιμο για τιμές της παραμέτρου  $f$  που δεν είναι μικρότερες από το λόγο προσέγγισης του αλγορίθμου. Η ανάλυση είναι αυστηρή και, το ενδιαφέρον είναι, ότι μπορεί να εφαρμοστεί και σε συμπληρωματικές εκδοχές του προβλήματος κάλυψης με σύνολα όπως η εξοικονόμηση χρωμάτων. Δίνει την ίδια εγγύηση προσέγγισης με τιμή  $193/240$  που οριακά βελτιώνει το προηγούμενο γνωστό κάτω φράγμα των Duh και Fűrer. Αποδεικνύουμε επίσης ότι, γενικά, μια φυσική κλάση αλγορίθμων τοπικής αναζήτησης δίνουν καλύτερα από  $1/2$ -προσεγγιστικά γεννητικά δάση αστέρων. Αυτά τα αποτελέσματα έχουν παρουσιαστεί στην εργασία [6].

### 3.1 Εισαγωγή

Μελετάμε το συνδυαστικό πρόβλημα αποσύνθεσης ενός γραφήματος σε ένα γεννητικό δάσος αστέρων που περιλαμβάνει όλες τις κορυφές του γραφήματος. Ένας *αστέρας* είναι ένα δένδρο με διάμετρο το πολύ δύο. Αποτελείται από μια κορυφή που χαρακτηρίζεται ως *κέντρο* και (πιθανώς) από *φύλλα*, δηλ., κορυφές που συνδέονται με ακμές με το κέντρο. Στόχος του προβλήματος που είναι γνωστό ως *πρόβλημα εύρεσης μέγιστου γεννητικού δάσους αστέρων* (SSF) είναι ο υπολογισμός μιας αποσύνθεσης αστέρα στην οποία ο αριθμός των φύλλων (που θεωρείται το κέρδος) μεγιστοποιείται. Ισοδύναμα, ο αριθμός των αστέρων πρέπει να ελαχιστοποιηθεί. Αυτό είναι ισοδύναμο με το πρόβλημα εύρεσης *ελάχιστου συνόλου κυριαρχίας* στο οποίο στόχος είναι ο υπολογισμός ενός ελάχιστου μεγέθους συνόλου κορυφών (το σύνολο κυριαρχίας) στο γράφημα τέτοιο ώστε κάθε κορυφή να είναι είτε μέλος του συνόλου κυριαρχίας ή γειτονική με κάποια κορυφή του συνόλου κυριαρχίας. Οι κορυφές στο σύνολο κυριαρχίας αντιστοιχούν σε κέντρα αστέρων με φύλλα τις κορυφές εκτός του συνόλου κυριαρχίας (βλ. Σχήμα 3.1).



Σχήμα 3.1: (i) Ένα γράφημα  $G = (V, E)$ , (ii) Ένα σύνολο κυριαρχίας ελάχιστου μεγέθους (3) στο  $G$  (κόκκινες κορυφές), (iii) Αποσύνθεση του  $G$  σε γεννητικό δάσος με 3 αστέρες.

Το πρόβλημα SSF έχει αρκετές εφαρμογές στο πρόβλημα παράταξης πολλαπλών γονιδιακών ακολουθιών [66], στη σύγκριση φυλογενετικών δένδρων [11], και στο πρόβλημα διαφοροποίησης στην αυτοκινητοβιομηχανία [1]. Ως ένα πρόβλημα συνδυαστικής βελτιστοποίησης, το πρόβλημα SSF έχει μελετηθεί στις εργασίες [23, 66]. Ο Nguyen και άλλοι [66] αποδεικνύουν ότι το πρόβλημα είναι δύσκολο να προσεγγιστεί (APX-hard) παρουσιάζοντας ένα

σαφές φράγμα μη προσεγγισιμότητας με τιμή  $259/260$  και έναν συνδυαστικό  $0.6$ -προσεγγιστικό αλγόριθμο. Βέλτιστοι αλγόριθμοι πολυωνυμικού χρόνου παρουσιάζονται για ειδικές κατηγορίες γραφημάτων όπως επίπεδα γραφήματα και δένδρα. Ο Chen και άλλοι [23] παρουσιάζουν έναν καλύτερο αλγόριθμο με λόγο προσέγγισης  $0.71$ . Η βασική ιδέα είναι η εκτέλεση και του αλγορίθμου της εργασίας [66], καθώς και ενός αλγορίθμου που βασίζεται σε γραμμικό προγραμματισμό και τυχαίο στρογγύλεμα και, μετά, η επιλογή της καλύτερης λύσης. Το γραμμικό πρόγραμμα με χαλαρωμένους περιορισμούς που χρησιμοποιείται στην εργασία [66] έχει ένα χάσμα ολοκλήρωσης ίσο με (το πολύ)  $3/4$ . Ενδιαφέρουσες γενικεύσεις περιλαμβάνουν εκδοχές του προβλήματος SSF με βάρη στις κορυφές και βάρη στις ακμές. Στις εργασίες [23, 66] παρουσιάζονται προσεγγιστικοί αλγόριθμοι και αποτελέσματα δυσκολίας προσέγγισης και για τα προβλήματα αυτά. Ισχυρότερα αποτελέσματα μη προσεγγισιμότητας για αυτά τα προβλήματα παρουσιάστηκαν πρόσφατα στην εργασία [21].

Το πρόβλημα εύρεσης ελάχιστου συνόλου κυριαρχίας αποτελεί ειδική περίπτωση του προβλήματος *κάλυψης με σύνολα*. Στο πρόβλημα κάλυψης με σύνολα, δίνεται μια συλλογή  $\mathcal{S}$  συνόλων σε ένα σύνολο  $U$  με  $n$  στοιχεία και στόχος είναι η επιλογή μιας ελάχιστου μεγέθους υπο-συλλογής  $\mathcal{T}$  συνόλων που περιέχουν όλα τα στοιχεία. Το πρόβλημα είναι γνωστό ότι είναι δύσκολο να προσεγγιστεί εντός ενός λογαριθμικού παράγοντα [38] ενώ ο άπληστος αλγόριθμος που αναδρομικά συμπεριλαμβάνει στην κάλυψη το σύνολο που περιέχει το μέγιστο πλήθος στοιχείων που δεν έχουν καλυφθεί ακόμα επιτυγχάνει ένα αντίστοιχο άνω φράγμα με τιμή  $H_n$  [52, 63]. Ο ίδιος αλγόριθμος έχει λόγο προσέγγισης με τιμή  $H_k$  όταν στη συλλογή ισχύει η κλεισιότητα ως προς τα υποσύνολα και κάθε σύνολο περιέχει το πολύ  $k$  στοιχεία. Αυτή η ειδική περίπτωση του προβλήματος είναι γνωστή ως πρόβλημα κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$ . Ελαφρώς καλύτερα αποτελέσματα είναι επίσης γνωστά για το πρόβλημα κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$  [4, 34, 43, 59]. Η πιο ενδιαφέρουσα ιδέα που δίνει αυτές τις βελτιώσεις είναι η *ημι-τοπική βελτίωση* [34] που εφαρμόζεται στο πρόβλημα κάλυψης 3-συνόλου. Η ημι-τοπική βελτιστοποίηση είναι ένας αλγόριθμος τοπικής αναζήτησης. Η βασική ιδέα στην οποία στηρίζεται είναι να ξεκινήσει κανείς με μία κενή κάλυψη 3-συνόλου και να την επαυξάνει πραγματοποιώντας τοπικές βελτιώσεις. Από τη στιγμή που τα (ξένα μεταξύ τους) σύνολα μεγέθους 3 επιλεχθούν, ο υπολογισμός του ελάχιστου αριθμού συνόλων μεγέθους 2 και 1 για την ολοκλήρωση της κάλυψης μπορεί να πραγματοποιηθεί σε πολυωνυμικό χρόνο μέσω του υπολογισμού ενός μέγιστου ταιριάσματος. Επομένως,

ένα ημι-τοπικό βήμα  $(s, t)$ -βελτίωσης για το πρόβλημα κάλυψης 3-συνόλου αποτελείται από (i) τη διαγραφή μέχρι το πολύ  $t$  συνόλων μεγέθους 3 από την τρέχουσα κάλυψη και (ii) την εισαγωγή μέχρι το πολύ  $s$  ξένων μεταξύ τους συνόλων μεγέθους 3 και του ελάχιστου αριθμού απαραίτητων συνόλων μικρότερου μεγέθους που ολοκληρώνουν την κάλυψη, έτσι ώστε ο αριθμός των συνόλων στην κάλυψη να μειώνεται. Στην περίπτωση ισοπαλιών, προτιμώνται καλύψεις με λιγότερα σύνολα μεγέθους 1. Η ανάλυση στην εργασία [34] δείχνει ότι η καλύτερη επιλογή των παραμέτρων  $(s, t)$  είναι  $(2, 1)$  και ότι, κάπως απρόσμενα, μεγαλύτερες τιμές αυτών των παραμέτρων δε δίνουν περαιτέρω βελτιώσεις.

Ως το συμπλήρωμα του προβλήματος εύρεσης ελάχιστου συνόλου κυριαρχίας, το πρόβλημα SSF αποτελεί ειδική περίπτωση του προβλήματος *συμπληρωματικού συνόλου κάλυψης* (με έννοια που εξηγείται λεπτομερώς στην Ενότητα 3.2). Στο συμπληρωματικό πρόβλημα κάλυψης με σύνολα, στόχος είναι η μεγιστοποίηση της ποσότητας  $n - |T|$ . Ο λόγος προσέγγισης μιας λύσης για το συμπληρωματικό πρόβλημα κάλυψης με σύνολα μπορεί να γίνει αντιληπτός σαν τη σύγκριση της ‘απόστασης’ μιας λύσης για το πρόβλημα κάλυψης με σύνολα από τη χειρότερη δυνατή λύση (δηλ., αυτή που χρησιμοποιεί  $n$  σύνολα για την κάλυψη των στοιχείων) με την απόσταση της βέλτιστης λύσης από τη χειρότερη δυνατή. Αυτό δίνει ένα εναλλακτικό μέτρο απόδοσης για την ανάλυση προσεγγιστικών αλγορίθμων. Τέτοια μέτρα έχουν ληφθεί υπόψη για πολλά προβλήματα συνδυαστικής βελτιστοποίησης στα πλαίσια διαφορετικών προσεγγιστικών αλγορίθμων [30, 31] ή  $z$ -προσεγγίσεων [46] (δες επίσης [8, 9]). Οι Duh και Fürer [34] μελετούν επίσης την εφαρμογή του αλγορίθμου τους για στιγμιότυπα του συμπληρωματικού προβλήματος κάλυψης με σύνολα στα οποία η συλλογή των συνόλων δε δίνεται ρητά. Μεταξύ αυτών, η πιο ενδιαφέρουσα περίπτωση είναι όταν τα στοιχεία αντιστοιχούν στις κορυφές ενός γραφήματος και τα σύνολα στη συλλογή αποτελούν όλα τα ανεξάρτητα σύνολα του γραφήματος. Στην περίπτωση αυτή, το πρόβλημα κάλυψης με σύνολα είναι ισοδύναμο με το πρόβλημα χρωματισμού γραφημάτων (δηλ., το πρόβλημα χρωματισμού των κορυφών ενός γραφήματος έτσι ώστε ποτέ δύο γειτονικές κορυφές να μη λαμβάνουν το ίδιο χρώμα και ο αριθμός των χρωμάτων που χρησιμοποιούνται να ελαχιστοποιείται). Το συμπληρωματικό πρόβλημα χρωματισμού γραφημάτων είναι επίσης γνωστό και με τον όρο πρόβλημα *εξοικονόμηση χρωμάτων*.

Εδώ, βασιζόμαστε στο γεγονός ότι το πρόβλημα SSF είναι ένα συμπληρωματικό πρόβλημα κάλυψης με σύνολα στο οποίο στη συλλογή των συνόλων ισχύει η κλειστότητα για τα υποσύνολα και χρησιμοποιούμε τον προ-

σεγγιστικό αλγόριθμο των Duh και Fürer [34] για να λύσουμε τα στιγμιότυπα του (Ενότητα 3.2). Επιτυγχάνουμε ένα λόγο προσέγγισης με τιμή  $193/240 \approx 0.804$  που βελτιώνει το προηγούμενο γνωστό φράγμα με τιμή 0.71 από την εργασία [23] και υπερκεράζει το χάσμα ολοκλήρωσης του χαλαρωμένου γραμμικού προγράμματος που χρησιμοποιείται στην εργασία [23]. Η ανάλυσή μας είναι αυστηρή, δεν επωφελείται από τη συγκεκριμένη δομή του του προβλήματος SSF, και ουσιαστικά ισχύει για το γενικότερο συμπληρωματικό πρόβλημα κάλυψης με σύνολα και εξοικονόμησης χρωμάτων επίσης (δες και Ενότητα 3.3). Αξίζει να παρατηρήσει κανείς ότι το αποτέλεσμα της εργασίας [34] αποτελεί ένα κάτω φράγμα με τιμή  $289/360 \approx 0.803$  για το λόγο προσέγγισης του ίδιου αλγόριθμου για γενικά στιγμιότυπα του προβλήματος εξοικονόμησης χρωμάτων (και του συμπληρωματικού προβλήματος κάλυψης με σύνολα). Αν και η βελτίωση που πετυχαίνουμε στο λόγο προσέγγισης είναι οριακή, η τεχνική απόδειξης παρουσιάζει μεγάλο ενδιαφέρον και, σαν ιδέα, θα μπορούσε να εφαρμοστεί και στα πλαίσια άλλων προβλημάτων. Η απόδειξη στην εργασία [34] βασίζεται σε μια λεπτομερή αποτίμηση της απόδοσης του αλγορίθμου μέσω ανάλυσης περιπτώσεων των διαφορετικών τρόπων κάλυψης των στοιχείων των συνόλων από τον αλγόριθμο στη βέλτιστη λύση. Μεταξύ των αρκετών διαφορετικών περιπτώσεων, μόνο μερικές παρουσιάζονται στην εργασία αυτή. Η δική μας απόδειξη είναι αρκετά διαφορετική σαν ιδέα, είναι απλούστερη, και δεν απαιτεί ανάλυση περιπτώσεων. Βασίζεται στην απόδειξη εφικτότητας ενός γραμμικού προγράμματος. Σημειώνουμε ότι η ανάλυση καθαρά συνδυαστικών αλγορίθμων που κάνουν χρήση γραμμικών προγραμμάτων των οποίων η αντικειμενική τιμή ‘αποκαλύπτει’ το λόγο προσέγγισης του υπό εξέταση αλγορίθμου (με διαφορετικό τρόπο από ότι στην παρούσα εργασία) έχει επίσης χρησιμοποιηθεί για το πρόβλημα κάλυψης με σύνολα με σύνολα μεγέθους το πολύ  $k$  [4], για διαχείριση οπτικών μηκών κύματος σε οπτικά δίκτυα [15], και για τοποθέτηση εγκαταστάσεων [51]. Ο ορισμός του γραμμικού προγράμματος ακολουθεί την ορολογία του προηγούμενου κεφαλαίου αλλά το γραμμικό πρόγραμμα που προκύπτει είναι σημαντικά μικρότερο και η ανάλυση διαφορετική. Συγκεκριμένα, αποδεικνύουμε ότι αν ένας αλγόριθμος επιτυγχάνει μία το πολύ  $f$ -προσεγγιστική λύση για κάποιο στιγμιότυπο, τότε ένα κατάλληλα ορισμένο γραμμικό πρόβλημα που έχει το  $f$  σα μια παράμετρο και οι περιορισμοί του συλλαμβάνουν τις ιδιότητες του στιγμιότυπου και τον τρόπο που αλγόριθμος εφαρμόζεται σε αυτό είναι εφικτό. Αυτό δηλώνεται σαν ένα *παραμετροποιημένο LP Λήμμα*. Επομένως, το κάτω φράγμα στο λόγο προσέγγισης προκύπτει αποδεικνύοντας ότι το αντίστοιχο γραμμικό πρόγραμμα δεν είναι εφικτό για τιμές του  $f$  μικρότερες από

193/240.

Επιπλέον, υποκινούμενοι από πολύ γνωστές ευριστικές τεχνικές για προβλήματα ομαδοποίησης συνόλου [43, 48], μελετάμε μια φυσική οικογένεια αλγορίθμων τοπικής αναζήτησης για το πρόβλημα SSF. Ένας τέτοιος αλγόριθμος ξεκινά από ένα αρχικό γεννητικό δάσος αστέρων και επιλαμβανόμενα πραγματοποιεί τοπικές βελτιώσεις μέχρις ότου να μην είναι περαιτέρω δυνατόν. Η λύση που προκύπτει είναι ένα τοπικό βέλτιστο για το συγκεκριμένο αλγόριθμο και η ερώτηση είναι αν τοπικά βέλτιστα αποτελούν αποδοτικές λύσεις και για το πρόβλημα SSF. Αποδεικνύουμε ότι αυτό δεν ισχύει για οποιοδήποτε αλγόριθμο τοπικής αναζήτησης που ανήκει στην οικογένεια και εκτελείται σε πολυωνυμικό χρόνο. Για την απόδειξη, κατασκευάζουμε σχεδόν  $1/2$ -προσεγγιστικά γεννητικά δάση αστέρων σε κατάλληλα ορισμένα γραφήματα που αποτελούν τοπικά βέλτιστα για τέτοιους αλγόριθμους (Ενότητα 3.5).

## 3.2 Περιγραφή του αλγορίθμου

Στην ενότητα αυτή, παρουσιάζουμε τον τρόπο με τον οποίο επωφελούμαστε τη σχέση του προβλήματος με το πρόβλημα κάλυψης με σύνολα προκειμένου να το επιλύσουμε χρησιμοποιώντας έναν πολύ γνωστό αλγόριθμο των Duh και Fürer [34]. Αρχικά μετατρέπουμε ένα στιγμιότυπο του προβλήματος SSF αποτελούμενο από ένα γράφημα  $G = (V, E)$  σε ένα στιγμιότυπο  $(V, S)$  του προβλήματος κάλυψης με σύνολα πάνω στις κορυφές. Η συλλογή  $S$  αποτελείται από *σύνολα αστέρων*. Ένα σύνολο αστέρων είναι ένα σύνολο κορυφών με έναν κοινό γείτονα (υποθέτοντας ότι μία κορυφή είναι γειτονική του εαυτού της). Διακρίνουμε μεταξύ δύο τύπων συνόλων αστέρα: A και B. Οι κορυφές ενός συνόλου αστέρα  $s$  τύπου A έχουν μια κορυφή του  $s$  σαν κοινό γείτονα. Όλα τα άλλα σύνολα αστέρων είναι τύπου B. Η συλλογή  $S$  που ορίζεται με τον τρόπο αυτό ισχύει η κλειστότητα ως προς τα υποσύνολα. Χρησιμοποιούμε τον όρο κάλυψη συνόλου αστέρα για να αναφερθούμε στο συγκεκριμένο στιγμιότυπο του προβλήματος κάλυψης με σύνολα και τον όρο κάλυψη συνόλου  $i$ -αστέρων όταν η συλλογή αποτελείται από σύνολα αστέρων μεγέθους το πολύ  $i$ .

Χρησιμοποιούμε τον αλγόριθμο των Duh και Fürer [34] για να λύσουμε το στιγμιότυπο του προβλήματος κάλυψης με σύνολα αστέρων που προκύπτει και παίρνουμε μια συλλογή ξένων μεταξύ τους συνόλων αστέρων που περιέχει όλες τις κορυφές του  $V$ . Αναφερόμαστε στα σύνολα αστέρων μεγέθους  $i$  με τον όρο σύνολα  $i$ -αστέρων. Ο αλγόριθμος αρχικά συμπεριλαμβάνει άπληστα

στη λύση ξένα μεταξύ τους σύνολα 6-αστέρων μέχρι μη μπορούν να συμπεριληφθούν άλλα σύνολα 6-αστέρων. Χρησιμοποιούμε τον όρο *μεγιστοτικές* (maximal) για να αναφερθούμε σε τέτοιες συλλογές ξένων μεταξύ τους συνόλων. Μετά τη φάση αυτή, οι κορυφές που δεν έχουν καλυφθεί και τα σύνολα αστέρων που περιέχουν μόνο τέτοιες κορυφές σχηματίζουν ένα στιγμιότυπο του προβλήματος κάλυψης με σύνολα 5-αστέρων. Τότε, ο αλγόριθμος εκτελεί μια περιορισμένη φάση για να επιλέξει μια μεγιστοτική συλλογή ξένων μεταξύ τους συνόλων 5-αστέρων και εγγυάται ότι ο αριθμός των συνόλων 1-αστέρα στην τελική λύση δεν είναι μεγαλύτερος από τον αριθμό των συνόλων 1-αστέρα στη βέλτιστη λύση για το στιγμιότυπο του προβλήματος κάλυψης με σύνολα 5-αστέρων στην αρχή αυτής της περιορισμένης φάσης. Στη συνέχεια, εφαρμόζεται μια παραπλήσια περιορισμένη φάση για σύνολα 4-αστέρων. Μετά τη φάση αυτή, ένα στιγμιότυπο του προβλήματος κάλυψης με σύνολα 3-αστέρων απομένει να λυθεί για να ολοκληρωθεί η κάλυψη συνόλου αστέρα. Ο αλγόριθμος εφαρμόζει για αυτό μια φάση ημι-τοπικής βελτιστοποίησης. Χρησιμοποιούμε τους ακέραιους 6, 5, 4, 3 για να αναφερθούμε στις φάσεις του αλγορίθμου σύμφωνα με το μέγεθος των συνόλων αστέρων που διαχειρίζονται. Συνολικά, ο αλγόριθμος περιγράφεται ως εξής.

**Φάση 6:** Επίλεξε μια μεγιστοτική συλλογή ξένων μεταξύ τους συνόλων 6-αστέρων.

**Φάση 5:** Επίλεξε μια μεγιστοτική συλλογή ξένων μεταξύ τους συνόλων 5-αστέρων έτσι ώστε η επιλογή αυτών των συνόλων αστέρα να μην αυξάνει τον αριθμό των συνόλων 1-αστέρα στην τελική λύση.

**Φάση 4:** Επίλεξε μια μεγιστοτική συλλογή ξένων μεταξύ τους συνόλων 4-αστέρων έτσι ώστε η επιλογή αυτών των συνόλων αστέρα να μην αυξάνει τον αριθμό των συνόλων 1-αστέρα στην τελική λύση.

**Φάση 3:** Εκτέλεσε τον αλγόριθμο ημι-τοπικής βελτιστοποίησης για το στιγμιότυπο του προβλήματος κάλυψης με σύνολα 3-αστέρων που απομένει.

Το αποτέλεσμα του αλγορίθμου είναι ένα σύνολο  $\mathcal{T} \subseteq \mathcal{S}$  ξένων μεταξύ τους συνόλων αστέρων. μετατρέπουμε τη λύση αυτή σε ένα γεννητικό δάσος αστέρων ως εξής. Αρχικά θεωρούμε σύνολα αστέρων του  $\mathcal{T}$  τύπου  $A$ . Για κάθε τέτοιο σύνολο  $s$ , θέτουμε τον κοινό γείτονα  $u$  των κορυφών του  $s$  σαν

κέντρο και συνδέουμε τις υπόλοιπες κορυφές κορυφές του  $s$  με την  $u$ . Στη συνέχεια, θεωρούμε τα σύνολα αστέρων τύπου B. Για κάθε τέτοιο σύνολο  $s$  του  $\mathcal{T}$ , επιλέγουμε έναν κοινό γείτονα  $u$  των κορυφών του  $s$ . Αν η  $u$  είναι ήδη κέντρο ενός άλλου αστέρα, απλά συνδέουμε τις κορυφές του  $s$  που δεν έχουν χρησιμοποιηθεί ακόμα σαν κέντρα με αυτή. Αν η  $u$  είναι ήδη φύλλο ενός άλλου αστέρα, την αφαιρούμε από τον αστέρα και συνδέουμε σε αυτή τις κορυφές του  $s$  που δεν έχουν χρησιμοποιηθεί σαν κέντρα ακόμα.

Από τη στιγμή που κάποια κορυφή έχει χρησιμοποιηθεί σαν κέντρο, παραμένει μέχρι το τέλος του αλγορίθμου και κάθε φύλλο συνδέεται με κάποιο κέντρο. Επειδή τα σύνολα αστέρων περιέχουν όλες τις κορυφές του  $V$ , κάθε κορυφή είναι είτε κέντρο είτε φύλλο ενός αστέρα και, επομένως, η λύση είναι πράγματι ένα γεννητικό δάσος αστέρων για το γράφημα  $G$ . Επιπλέον, κάθε φορά που ορίζουμε ένα σύνολο αστέρα, αυξάνουμε τον αριθμό των κέντρων κατά το πολύ 1. Επομένως, ο συνολικός αριθμός των φύλλων είναι τουλάχιστον  $|V| - |\mathcal{T}|$ .

### 3.3 Ανάλυση

Το βασικό μας επιχείρημα για την ανάλυση του αλγορίθμου περιγράφεται ως εξής. Αρχικά αποδεικνύουμε ότι αν υπάρχει ένα στιγμιότυπο για το οποίο ο αλγόριθμος υπολογίζει μια το πολύ  $f$ -προσεγγιστική λύση, τότε ένα κατάλληλα ορισμένο γραμμικό πρόγραμμα  $LP(f)$  είναι εφικτό. Στη συνέχεια, δείχνουμε ότι το  $LP(f)$  δεν είναι εφικτό για  $f < 193/240$ , που σημαίνει ότι ο λόγος προσέγγισης του αλγορίθμου είναι τουλάχιστον  $193/240 \approx 0.804$ . Θεωρούμε μόνο μη τριμμένα στιγμιότυπα του προβλήματος στα οποία το γράφημα εισόδου έχει τουλάχιστον μία ακμή αφού διαφορετικά κάθε αλγόριθμος είναι βέλτιστος.

#### 3.3.1 Ένα παραμετροποιημένο γραμμικό πρόγραμμα

Θεωρούμε ένα μη τριμμένο στιγμιότυπο  $I$  του προβλήματος SSF που αποτελείται από ένα γράφημα  $G = (V, E)$  για οποίο ο αλγόριθμος υπολογίζει μια το πολύ  $f$ -προσεγγιστική λύση. Έστω  $(V, \mathcal{S})$  το αντίστοιχο στιγμιότυπο του προβλήματος κάλυψης με σύνολα αστέρα. Για  $i = 5, 4, 3$ , συμβολίζουμε με  $(V_i, \mathcal{S}_i)$  το στιγμιότυπο του προβλήματος κάλυψης με σύνολα  $i$ -αστέρα που πρέπει να λυθεί πριν την έναρξη της φάσης  $i$ . Εδώ, τα σύνολα  $V_i$  περιέχουν τις κορυφές του  $V$  που δεν έχουν καλυφθεί σε προηγούμενες φάσεις και τα



$S_i$  περιέχουν τα σύνολα αστέρων του  $\mathcal{S}$  που αποτελούνται μόνο από κορυφές του  $V_i$ . Τα  $S_i$  περιέχουν σύνολα αστέρων της αρχικής συλλογής μεγέθους το πολύ  $i$ . Συμβολίζουμε με  $\mathcal{O}_i$  μία βέλτιστη λύση για ένα στιγμιότυπο  $(V_i, S_i)$ . Συμβολίζουμε τη βέλτιστη λύση του  $(V, \mathcal{S})$  με  $\mathcal{O}$ . Επειδή στο  $\mathcal{S}$  ισχύει η κλειστότητα ως προς τα υποσύνολα, χωρίς απώλεια της γενικότητας, μπορούμε να υποθέσουμε ότι το σύνολο  $\mathcal{O}_i$  περιέχει ξένα μεταξύ τους σύνολα. Επιπλέον, είναι σαφές πλέον ότι  $|\mathcal{O}_{i-1}| \leq |\mathcal{O}_i|$  για  $i = 3, 4, 5$  και  $|\mathcal{O}_i| \leq |\mathcal{O}|$ .

Θέτουμε  $\theta = \frac{|V|}{|V|-1}$ . Συμβολίζουμε με  $T$  το λόγο  $|V|/|\mathcal{O}|$ . Επειδή το στιγμιότυπο δεν είναι τετριμμένο, υπάρχει μια βέλτιστη κάλυψη αστέρα με το πολύ  $|V| - 1$  σύνολα αστέρα (δηλ.,  $|\mathcal{O}| \leq |V| - 1$ ). Επομένως,

$$T \geq \theta. \quad (3.1)$$

Για τη φάση  $i$  του αλγορίθμου με  $i = 3, 4, 5$ , συμβολίζουμε με  $a_{i,j}$  το λόγο του αριθμού των συνόλων  $j$ -αστέρων στο  $\mathcal{O}_i$  προς τον αριθμό  $|\mathcal{O}|$  των συνόλων στη βέλτιστη λύση για  $(V, \mathcal{S})$ . Επειδή  $|\mathcal{O}_i| \leq |\mathcal{O}|$  και  $|\mathcal{O}_i| = |\mathcal{O}| \sum_{j=1}^i a_{i,j}$ , έχουμε ότι

$$\sum_{j=1}^i a_{i,j} \leq 1. \quad (3.2)$$

Επίσης, παρατηρούμε ότι  $|V| = T|\mathcal{O}|$  και  $|V_5| = |\mathcal{O}| \sum_{j=1}^5 j a_{5,j}$ . Επειδή  $V_5 \subseteq V$ , έχουμε

$$T \geq \sum_{j=1}^5 j a_{5,j}. \quad (3.3)$$

Κατά τη φάση  $i$  με  $i = 3, 4, 5$ , ο αλγόριθμος επιλέγει ένα μέγιστο σύνολο από σύνολα  $i$ -αστέρων. Αυτό σημαίνει ότι κάθε σύνολο  $i$ -αστέρα που επιλέγεται τέμνει κάποιο από τα σύνολα  $i$ -αστέρα της βέλτιστης κάλυψης συνόλων αστέρα για το στιγμιότυπο  $(V_i, S_i)$  και ενδεχομένως τέμνεται με το πολύ  $i$  τέτοια σύνολα  $i$ -αστέρων. Επειδή υπάρχουν  $a_{i,i}|\mathcal{O}|$  τέτοια σύνολα αστέρων στη βέλτιστη κάλυψη για το  $(V_i, S_i)$ , ο αριθμός  $|V_i \setminus V_{i-1}|/i$  των συνόλων  $i$ -αστέρα που επιλέχθηκαν κατά τη φάση  $i$  είναι τουλάχιστον  $a_{i,i}|\mathcal{O}|/i$ . Ισοδύναμα,  $|V_i \setminus V_{i-1}| \geq a_{i,i}|\mathcal{O}|$ . Επειδή  $|V_i \setminus V_{i-1}| = \left( \sum_{j=1}^i j a_{i,j} - \sum_{j=1}^{i-1} j a_{i-1,j} \right) |\mathcal{O}|$ , έχουμε ότι

$$\sum_{j=1}^i j a_{i,j} - \sum_{j=1}^{i-1} j a_{i-1,j} \geq a_{i,i}. \quad (3.4)$$

Η φάση 3 προσθέτει αρκετούς επιπλέον περιορισμούς. Συμβολίζουμε με  $b_3, b_2, b_1$  τον αριθμό των συνόλων 3-, 2-, και 1-αστέρων που υπολογίζονται από τη φάση ημι-τοπικής βελτιστοποίησης διαιρεμένο με  $|\mathcal{O}|$ . Χρησιμοποιούμε το κεντρικό αποτέλεσμα της ανάλυσης των Duh και Furer [34] εκφρασμένο με βάση τους δικούς μας συμβολισμούς.

**Θεώρημα 9** (Duh και Furer[34]).  $b_1 \leq \alpha_{3,1}$  και  $b_2 + b_1 \leq \alpha_{3,3} + \alpha_{3,2} + \alpha_{3,1}$ .

Επιπλέον, η περιορισμένη φάση  $i$  (με  $i = 4, 5$ ) θέτει τον περιορισμό ότι ο αριθμός των 1-συνόλων στην τελική λύση δεν πρέπει να αυξάνεται σε σχέση με τον αριθμό των 1-συνόλων στη βέλτιστη κάλυψη συνόλου αστέρα για το  $(V_i, S_i)$ . Λαμβάνοντας υπόψη και την πρώτη ανισότητα του Θεωρήματος 9, έχουμε

$$b_1 \leq a_{i,1}, \text{ για } i = 3, 4, 5. \quad (3.5)$$

Έως τώρα, έχουμε εκφράσει όλες τις ιδιότητες του στιγμιότυπου  $(V, S)$  καθώς και τη συμπεριφορά του αλγορίθμου για αυτό, εκτός από το γεγονός ότι η κάλυψη συνόλου αστέρων που επιτυγχάνεται συνεπάγεται μία το πολύ  $f$ -προσεγγιστική λύση για το αντίστοιχο στιγμιότυπο του προβλήματος SSF στιγμιότυπο. Εκφράζουμε το κέρδος του αλγορίθμου ως προς τις μεταβλητές μας ως εξής. Συμβολίζουμε με  $t_i$  τον αριθμό των συνόλων αστέρων που υπολογίζονται κατά τη φάση  $i$ . Για τη φάση 6, έχουμε

$$t_6 = \frac{1}{6}|V \setminus V_5| = \frac{1}{6} \left( T - \sum_{j=1}^5 j a_{5,j} \right) |\mathcal{O}|. \quad (3.6)$$

Για τη φάση  $i$  με  $i = 4, 5$ , έχουμε

$$t_i = \frac{1}{i}|V_i \setminus V_{i-1}| = \frac{1}{i} \left( \sum_{j=1}^i j a_{i,j} - \sum_{j=1}^{i-1} j a_{i-1,j} \right) |\mathcal{O}|. \quad (3.7)$$

Επίσης, για τη φάση ημι-τοπικής βελτιστοποίησης, έχουμε

$$\begin{aligned} t_3 &= (b_3 + b_2 + b_1) |\mathcal{O}| \\ &= \left( \frac{b_1}{3} + \frac{b_2 + b_1}{3} + \frac{3b_3 + 2b_2 + b_1}{3} \right) |\mathcal{O}| \\ &\leq \left( \frac{b_1}{3} + \frac{a_{3,3} + a_{3,2} + a_{3,1}}{3} + \frac{3a_{3,3} + 2a_{3,2} + a_{3,1}}{3} \right) |\mathcal{O}| \\ &= \left( \frac{1}{3}b_1 + \frac{4}{3}a_{3,3} + a_{3,2} + \frac{2}{3}a_{3,1} \right) |\mathcal{O}|. \end{aligned} \quad (3.8)$$

Η ανισότητα προκύπτει από το Θεώρημα 9 και επειδή  $(3b_3 + 2b_2 + b_1)|\mathcal{O}| = (3a_{3,3} + 2a_{3,2} + a_{3,1})|\mathcal{O}| = |V_3|$ .

Από τη συζήτηση στην Ενότητα 3.2, η λύση που επιτυγχάνεται από τον αλγόριθμο για το στιγμιότυπο  $I$  του προβλήματος SSF έχει κέρδος  $ALG(I)$  τουλάχιστον  $|V| - \sum_{i=3}^6 t_i$ . Έστω ένα βέλτιστο γεννητικό δάσος αστέρων για το γράφημα  $G$  που αντιστοιχεί στο στιγμιότυπο  $I$  και έστω  $OPT(I)$  το κέρδος του. Αυτό φυσικά αντιστοιχεί σε μία κάλυψη συνόλου αστέρα  $\mathcal{O}'$  για  $(V, \mathcal{S})$  που αποτελείται από ξένα μεταξύ τους υποσύνολα αστέρα τύπου A. Προφανώς,  $OPT(I) = |V| - |\mathcal{O}'|$  και  $|\mathcal{O}| \leq |\mathcal{O}'|$ . Επομένως,  $OPT(I) \leq |V| - |\mathcal{O}|$ . Επειδή η λύση που επιτυγχάνεται από τον αλγόριθμο είναι το πολύ  $f$ -προσεγγιστική (δηλ.,  $ALG(I) \leq f \cdot OPT(I)$ ), έχουμε

$$|V| - \sum_{i=3}^6 t_i \leq f(|V| - |\mathcal{O}|),$$

και, ισοδύναμα,

$$(f-1)|V| + \sum_{i=3}^6 t_i \geq f|\mathcal{O}|. \quad (3.9)$$

Χρησιμοποιούμε τις (3.6), (3.7), και (3.8) για να φράξουμε από πάνω το αριστερό μέρος της ανισότητας (3.9). Έχουμε

$$\begin{aligned} & (f-1)|V| + \sum_{i=3}^6 t_i \\ & \leq (f-1)T|\mathcal{O}| + \frac{1}{6} \left( T - \sum_{j=1}^5 ja_{5,j} \right) |\mathcal{O}| + \frac{1}{5} \left( \sum_{j=1}^5 ja_{5,j} - \sum_{j=1}^4 ja_{4,j} \right) |\mathcal{O}| \\ & \quad + \frac{1}{4} \left( \sum_{j=1}^4 ja_{4,j} - \sum_{j=1}^3 ja_{3,j} \right) |\mathcal{O}| + \left( \frac{1}{3}b_1 + \frac{4}{3}a_{3,3} + a_{3,2} + \frac{2}{3}a_{3,1} \right) |\mathcal{O}| \\ & = \left( \left( f - \frac{5}{6} \right) T + \frac{1}{6}a_{5,5} + \frac{2}{15}a_{5,4} + \frac{1}{10}a_{5,3} + \frac{1}{15}a_{5,2} + \frac{1}{30}a_{5,1} + \frac{1}{5}a_{4,4} \right. \\ & \quad \left. + \frac{3}{20}a_{4,3} + \frac{1}{10}a_{4,2} + \frac{1}{20}a_{4,1} + \frac{7}{12}a_{3,3} + \frac{1}{2}a_{3,2} + \frac{5}{12}a_{3,1} + \frac{1}{3}b_1 \right) |\mathcal{O}|. \end{aligned} \quad (3.10)$$

Από (3.9) και (3.10), έχουμε

$$\begin{aligned} & (f - \frac{5}{6})T + \frac{1}{6}a_{5,5} + \frac{2}{15}a_{5,4} + \frac{1}{10}a_{5,3} + \frac{1}{15}a_{5,2} + \frac{1}{30}a_{5,1} + \frac{1}{5}a_{4,4} \\ & + \frac{3}{20}a_{4,3} + \frac{1}{10}a_{4,2} + \frac{1}{20}a_{4,1} + \frac{7}{12}a_{3,3} + \frac{1}{2}a_{3,2} + \frac{5}{12}a_{3,1} + \frac{1}{3}b_1 \geq f. \end{aligned} \quad (3.11)$$

Εκφράζοντας τις ανισότητες (3.1)-(3.5) και (3.11) στη βασική μορφή, έχουμε το παραμετροποιημένο LP λήμμα.

**Λήμμα 10.** *Αν υπάρχει ένα στιγμιότυπο  $I$  του προβλήματος SSF για το οποίο ο αλγόριθμος υπολογίζει μια λύση με κέρδος  $ALG(I) \leq f \cdot OPT(I)$  για κάποιο  $f \in [0, 1]$ , τότε το ακόλουθο γραμμικό πρόβλημα  $LP(f)$  έχει μία εφικτή λύση για κάποιο  $\theta > 1$ .*

$$\begin{aligned} & T \geq \theta \\ & - \sum_{j=1}^i a_{i,j} \geq -1, \quad \text{για } i = 3, 4, 5 \\ & T - \sum_{j=1}^5 ja_{5,j} \geq 0 \\ & (i-1)a_{i,i} + \sum_{j=1}^{i-1} ja_{i,j} - \sum_{j=1}^{i-1} ja_{i-1,j} \geq 0, \quad \text{για } i = 4, 5 \\ & a_{i,1} - b_1 \geq 0, \quad \text{για } i = 3, 4, 5 \\ & (f - \frac{5}{6})T + \frac{1}{6}a_{5,5} + \frac{2}{15}a_{5,4} + \frac{1}{10}a_{5,3} + \frac{1}{15}a_{5,2} + \frac{1}{30}a_{5,1} + \frac{1}{5}a_{4,4} \\ & + \frac{3}{20}a_{4,3} + \frac{1}{10}a_{4,2} + \frac{1}{20}a_{4,1} + \frac{7}{12}a_{3,3} + \frac{1}{2}a_{3,2} + \frac{5}{12}a_{3,1} + \frac{1}{3}b_1 \geq f \\ & a_{i,j} \geq 0, \quad \text{για } i = 3, 4, 5 \text{ και } j = 1, \dots, i \\ & b_1 \geq 0 \end{aligned}$$

### 3.3.2 Απόδειξη του φράγματος για το λόγο προσέγγισης

Η απόδειξη του φράγματος προσέγγισης βασίζεται στο ακόλουθο λήμμα.

**Λήμμα 11.** *Για κάθε  $f < 193/240$ , το γραμμικό πρόγραμμα  $LP(f)$  δεν έχει εφικτή λύση.*

*Απόδειξη.* Μπορούμε να υποθέσουμε ότι το  $LP(f)$  είναι ένα γραμμικό πρόγραμμα ελαχιστοποίησης με αντικείμενο 0. Από δυσκότητα, αν ήταν εφικτό, η βέλτιστη αντικειμενική τιμή του δυικού γραμμικού προγράμματος μεγιστοποίησης θα ήταν επίσης 0. Δείχνουμε ότι κάτι τέτοιο δε συμβαίνει και ότι το δυικό γραμμικό πρόγραμμα έχει μία λύση με αυστηρά θετική αντικειμενική τιμή. Έτσι προκύπτει το λήμμα.

Στο δυικό γραμμικό πρόγραμμα, χρησιμοποιούμε τις έντεκα μεταβλητές  $\eta, \beta_3, \beta_4, \beta_5, \gamma_4, \gamma_5, \gamma_6, \delta_3, \delta_4, \delta_5$  και  $\zeta$  που αντιστοιχούν στους περιορισμούς του γραμμικού προγράμματος  $LP(f)$ . Η μεταβλητή  $\eta$  αντιστοιχεί στον πρώτο περιορισμό του  $LP(f)$ , οι  $\beta_i$  αντιστοιχούν στο δεύτερο σύνολο περιορισμών, η  $\gamma_6$  αντιστοιχεί στον τρίτο περιορισμό, οι  $\gamma_4$  και  $\gamma_5$  αντιστοιχούν στο τέταρτο σύνολο περιορισμών, οι  $\delta_i$  αντιστοιχούν στο πέμπτο σύνολο περιορισμών, και η  $\zeta$  αντιστοιχεί στον τελευταίο περιορισμό. Επομένως, το δυικό γραμμικό πρόγραμμα για το  $LP(f)$  είναι

$$\begin{aligned}
 & \text{maximize} && \theta\eta - \beta_3 - \beta_4 - \beta_5 + f\zeta \\
 & \text{subject to} && \eta + \gamma_6 + (f - 5/6)\zeta \leq 0 \\
 & && -\delta_3 - \delta_4 - \delta_5 + \zeta/3 \leq 0 \\
 & && -\beta_3 - \gamma_4 + \delta_3 + 5\zeta/12 \leq 0 \\
 & && -\beta_3 - 2\gamma_4 + \zeta/2 \leq 0 \\
 & && -\beta_3 - 3\gamma_4 + 7\zeta/12 \leq 0 \\
 & && -\beta_4 + \gamma_4 - \gamma_5 + \delta_4 + \zeta/20 \leq 0 \\
 & && -\beta_4 + 2\gamma_4 - 2\gamma_5 + \zeta/10 \leq 0 \\
 & && -\beta_4 + 3\gamma_4 - 3\gamma_5 + 3\zeta/20 \leq 0 \\
 & && -\beta_4 + 3\gamma_4 - 4\gamma_5 + \zeta/5 \leq 0 \\
 & && -\beta_5 + \gamma_5 - \gamma_6 + \delta_5 + \zeta/30 \leq 0 \\
 & && -\beta_5 + 2\gamma_5 - 2\gamma_6 + \zeta/15 \leq 0 \\
 & && -\beta_5 + 3\gamma_5 - 3\gamma_6 + \zeta/10 \leq 0 \\
 & && -\beta_5 + 4\gamma_5 - 4\gamma_6 + 2\zeta/15 \leq 0 \\
 & && -\beta_5 + 4\gamma_5 - 5\gamma_6 + \zeta/6 \leq 0 \\
 & && \beta_i, \delta_i \geq 0, \quad \text{για } i = 3, 4, 5 \\
 & && \gamma_i \geq 0, \quad \text{για } i = 4, 5, 6 \\
 & && \zeta, \eta \geq 0
 \end{aligned}$$

Η λύση  $\eta = 193/240 - f$ ,  $\beta_3 = 39/72$ ,  $\beta_4 = 11/72$ ,  $\beta_5 = 79/720$ ,  $\gamma_4 = 1/72$ ,  $\gamma_5 = 1/45$ ,  $\gamma_6 = 7/240$ ,  $\delta_3 = 5/36$ ,  $\delta_4 = 1/9$ ,  $\delta_5 = 1/12$ , και  $\zeta = 1$  ικανοποιεί όλους τους περιορισμούς. Παρατηρούμε ότι  $\eta - \beta_3 - \beta_4 - \beta_5 + f\zeta = 0$  και, επομένως, η αντικειμενική τιμή είναι  $(\theta - 1)\eta = (\theta - 1)(193/240 - f) > 0$  και προκύπτει το λήμμα.  $\square$

**Θεώρημα 12.** Ο λόγος προσέγγισης του αλγορίθμου είναι τουλάχιστον  $193/240$ .

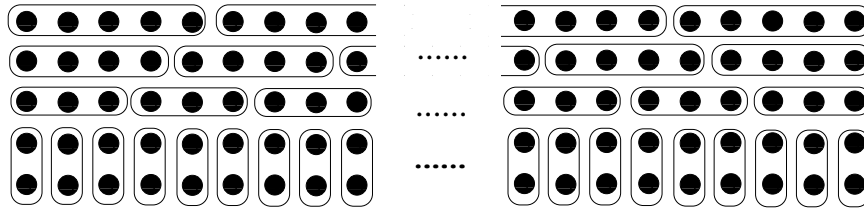
*Απόδειξη.* Από τα Λήμματα 10 και 11, έχουμε ότι για κάθε  $f < 193/240$  και για κάθε στιγμιότυπο  $I$  του προβλήματος SSF, ο αλγόριθμος υπολογίζει μία λύση με κέρδος  $ALG(I) > f \cdot OPT(I)$ . Επομένως, ο λόγος προσέγγισής του είναι τουλάχιστον  $193/240$ .  $\square$

### 3.4 Εξοικονόμηση χρωμάτων

Παρατηρούμε πως ούτε ο αλγόριθμος ούτε η ανάλυση δε χρησιμοποιούν το γεγονός ότι τα σύνολα αστέρων ουσιαστικά αντιστοιχούν σε αστέρες στο γράφημα. Επομένως, ο αλγόριθμος μπορεί να εφαρμοστεί και στο γενικότερο συμπληρωματικό πρόβλημα κάλυψης με σύνολα στο οποίο τα σύνολα δίνονται ρητά. Επίσης, αφού ο αλγόριθμος λαμβάνει υπόψη σύνολα σταθερού μεγέθους, μπορεί να εφαρμοστεί και για το πρόβλημα εξοικονόμησης χρωμάτων. Αυτό που απαιτείται είναι η αντικατάσταση του όρου ‘σύνολο αστέρα’ από τον όρο ‘ανεξάρτητο σύνολο’. Αν και ο υπολογισμός ενός ανεξάρτητου συνόλου δεδομένου μεγέθους σε ένα γράφημα είναι ένα κλασικό δύσκολο (NP-hard) πρόβλημα, ο υπολογισμός ανεξάρτητων συνόλων μεγέθους το πολύ  $\delta$  που απαιτεί ο αλγόριθμος μπορεί να πραγματοποιηθεί εύκολα σε πολυωνυμικό χρόνο.

Στο σχήμα 3.2, παρουσιάζουμε ένα παράδειγμα που δείχνει ότι, τουλάχιστον για το πρόβλημα εξοικονόμησης χρωμάτων, η ανάλυσή μας στην Ενότητα 3.3 είναι αυστηρή. Ανάλογες αλλά πιο περίπλοκες κατασκευές ισχύουν και για το πρόβλημα SSF. Το στιγμιότυπο εισόδου αποτελείται από ένα γράφημα με 300 κορυφές. Τα μαύρα σφαιρίδια αντιστοιχούν σε κορυφές. Οι κορυφές που ανήκουν στο ίδιο παραλληλόγραμμο με στρογγυλεμένες γωνίες ή στην ίδια στήλη αποτελούν ένα ανεξάρτητο σύνολο. Η βέλτιστη λύση αποτελείται από τα 60 ανεξάρτητα σύνολα μεγέθους 5 που αντιστοιχούν σε κάθε στήλη σφαιριδίων ενώ ο αλγόριθμος υπολογίζει τα ανεξάρτητα σύνολα που αντιστοιχούν στα παραλληλόγραμμο με τις στρογγυλεμένες γωνίες. Επομένως, το γράφημα περιέχει μια ακμή μεταξύ οποιουδήποτε ζεύγους κορυφών που δεν

ανήκουν στο ίδιο ανεξάρτητο σύνολο με αυτές που περιγράφηκαν προηγουμένως. Σημειώνουμε ότι μόνο 18 από τις 60 στήλες σφαιριδίων φαίνονται. Ο συνολικός αριθμός συνόλων που χρησιμοποιείται είναι 107 και δίνει μια λύση με κέρδος 193 για το πρόβλημα εξοικονόμησης χρωμάτων ενώ το βέλτιστο κέρδος είναι 240.



Σχήμα 3.2: Ένα αυστηρό κάτω φράγμα για το πρόβλημα εξοικονόμησης χρωμάτων.

### 3.5 Ένα κάτω φράγμα για αλγόριθμους τοπικής αναζήτησης

Στην ενότητα αυτή, μελετάμε μια φυσική οικογένεια αλγορίθμων τοπικής αναζήτησης για το πρόβλημα SSF. Θεωρούμε λύσεις στιγμιοτύπων του προβλήματος SSF αναθέσεις των τιμών 0 και 1 στις κορυφές του γραφήματος εισόδου όπου 0 ή 1 σε κάποια κορυφή σημαίνει ότι η κορυφή είναι κέντρο ή φύλλο, αντίστοιχα. Μια τέτοια ανάθεση αντιστοιχεί σε μια εφικτή λύση για το πρόβλημα SSF όταν κάθε κορυφή στην οποία έχει ανατεθεί η τιμή 1 είναι γειτονική με τουλάχιστον μία κορυφή στην οποία έχει ανατεθεί η τιμή 0. Το κέρδος μιας εφικτής ανάθεσης είναι τότε ο αριθμός των κορυφών στις οποίες έχει ανατεθεί η τιμή 1.

Η τοπική αναζήτηση μπορεί να χρησιμοποιηθεί ως εξής. Ξεκινώντας από οποιαδήποτε εφικτή ανάθεση (δηλ., με τιμή 0 σε κάθε κορυφή), ένας αλγόριθμος τοπικής αναζήτησης πραγματοποιεί αναδρομικά  $k$ -αλλαγές (για κάποια ακέραια σταθερά  $k$ ) όσο αυτό είναι δυνατόν. Η πραγματοποίηση μιας  $k$ -αλλαγής σημαίνει η εναλλαγή των αναθέσεων  $t \leq k$  κορυφών οι οποίες είχαν αρχικά την τιμή 1 σε 0 και η ανάθεση σε  $t + 1$  κορυφές που αρχικά είχαν την τιμή 0 της τιμής 1. Ο αλγόριθμος τερματίζει όταν έχει επιτευχθεί μια τοπικά βέλτιστη ανάθεση, δηλ., μία από την οποία καμία  $k$ -αλλαγή δεν είναι

δυνατή. Αφού ο όρος  $k$  είναι σταθερά και το κέρδος αυξάνει κατά 1 σε κάθε βήμα, ο αλγόριθμος τερματίζει σε πολυωνυμικό χρόνο. Ο αλγόριθμος είναι αποδοτικός αν όλα τα τοπικά βέλτιστα έχουν υψηλό κέρδος (σε σύγκριση με τη βέλτιστη εφικτή ανάθεση).

Ο αλγόριθμος τοπικής αναζήτησης που πραγματοποιεί 0-αλλαγές μπορεί να έχει πολύ κακό λόγο προσέγγισης. Πράγματι, θεωρούμε το στιγμιότυπο που αποτελείται από δύο κορυφές  $u$  και  $v$  στις οποίες έχει ανατεθεί η τιμή 1 και  $n - 2$  κορυφές  $u_1, \dots, u_{n-2}$  που συνδέονται και με τη  $u$  και με τη  $v$  και σε όλες έχει ανατεθεί η τιμή 0. Η ανάθεση αυτή έχει κέρδος 2 και είναι ένα τοπικό βέλτιστο αφού καμία 0-αλλαγή δεν είναι δυνατή, ενώ η λύση που αναθέτει την τιμή 1 στις κορυφές  $u_1, \dots, u_{n-2}$  και την τιμή 0 στις κορυφές  $u$  και  $v$  είναι εφικτή και έχει κέρδος  $n - 2$ . Επιπλέον, μπορούμε να αποδείξουμε ότι ο αλγόριθμος τοπικής αναζήτησης που πραγματοποιεί 1-αλλαγές πάντα υπολογίζει μία  $1/2$ -προσεγγιστική ανάθεση. Αυτό θα μπορούσε να υποδεικνύει ότι καλύτερα φράγματα θα μπορούσαν να επιτευχθούν μελετώντας  $k$ -αλλαγές με μεγαλύτερες σταθερές τιμές για το  $k$ . Δυστυχώς, αυτό δεν ισχύει όπως φαίνεται στο ακόλουθο θεώρημα.

**Θεώρημα 13.** Για κάθε ακέραιο  $k \geq 0$  και κάθε  $\epsilon \in \left(0, \frac{1}{2(k+2)}\right]$ , ο αλγόριθμος τοπικής αναζήτησης που πραγματοποιεί  $k$ -αλλαγές έχει μία  $(1/2 + \epsilon)$ -προσεγγιστική λύση σαν τοπικό βέλτιστο.

*Απόδειξη.* Η απόδειξη χρησιμοποιεί ένα αποτέλεσμα των Erdős και Sachs [36] σύμφωνα με το οποίο, για κάθε ακέραιο  $d, g > 0$ , υπάρχει ένα  $d$ -κανονικό γράφημα με περιφέρεια (girth) τουλάχιστον  $g$ .

Ξεκινάμε από ένα  $d$ -κανονικό γράφημα  $G$  με περιφέρεια τουλάχιστον  $g$ , όπου  $d = 2\lceil 1/2\epsilon \rceil$  και  $g = k + 2$ . Δεδομένου του  $G$ , θεωρούμε το γράφημα  $G'$  που προκύπτει αν αντικαταστήσουμε κάθε ακμή  $(u, v)$  του  $G$  με ένα μονοπάτι μεγέθους 3 (που από εδώ και στο εξής καλούμε ένα 3-μονοπάτι)  $\langle u, Z_{uv}^1, Z_{uv}^2, v \rangle$ . Επομένως, το γράφημα  $G$  έχει δύο παραπάνω κορυφές  $Z_{uv}^1$  και  $Z_{uv}^2$  για κάθε ακμή  $(u, v)$  στο  $G$ . Αναφερόμαστε σε αυτές τις κορυφές με τον όρο ακμές-κορυφές.

Ορίζουμε μια ανάθεση στις κορυφές του  $G'$  που έχει (ασυμπτωτικά) μισό κέρδος από το βέλτιστο και η οποία δε μπορεί να βελτιωθεί από μια  $k$ -αλλαγή. Για να κατασκευάσουμε την ανάθεση αυτή, χρησιμοποιούμε το γεγονός ότι ένα συνεκτικό γράφημα έχει ένα κύκλωμα Euler όταν όλες οι κορυφές του έχουν άρτιο βαθμό. Θεωρούμε ένα τέτοιο κατευθυνόμενο κύκλωμα Euler στο  $G'$  και αναθέτουμε τιμή 1 στην πρώτη κορυφή κάθε 3-μονοπατιού από



το οποίο διέρχεται το κύκλωμα Euler και τιμή 0 στην άλλη κορυφή του 3-μονοπατιού. Επιπλέον, αναθέτουμε την τιμή 1 σε κάθε κορυφή που δεν είναι ακμή-κορυφή. Αυτή είναι μια εφικτή ανάθεση αφού κάθε ακμή-κορυφή με τιμή 1 συνδέεται με την άλλη ακμή-κορυφή του ίδιου 3-μονοπατιού που έχει τιμή 0, και κάθε κορυφή που δεν είναι ακμή-κορυφή έχει ακριβώς  $d/2 = \lceil 1/2\epsilon \rceil \geq k + 2$  γείτονες με τιμή 0.

Αποδεικνύουμε με απαγωγή σε άτοπο. Υποθέτουμε ότι η παραπάνω κακή ανάθεση δεν είναι ένα τοπικό βέλτιστο και θεωρούμε μια άλλη γειτονική εφικτή ανάθεση (που επιτυγχάνεται με εφαρμογή μιας  $k$ -αλλαγής στην αρχική). Ειδικότερα, υποθέτουμε ότι η βελτιωμένη ανάθεση επιτυγχάνεται αν αναθέσουμε την τιμή 1 σε  $t + 1$  ακμές-κορυφές που έχουν τιμή 0 αρχικά, καθώς και αν αλλάξουμε την τιμή  $s \leq t$  ακμών-κορυφών και  $t - s$  κορυφών που δεν είναι ακμές-κορυφές από 1 σε 0.

Θεωρούμε μια ενδιάμεση ανάθεση, στην οποία έχουν εφαρμοστεί μόνο οι αλλαγές στις τιμές των ακμών-κορυφών. Παρατηρούμε ότι οι περιορισμοί σχετικά με τις κορυφές που δεν είναι ακμές-κορυφές δεν παραβιάζονται αφού κάθε τέτοια κορυφή αρχικά είχε τουλάχιστον  $k+2$  γείτονες με τιμή 0 και εφαρμόστηκαν το πολύ  $k + 1$  αλλαγές από 0 σε 1. Σχετικά με τις ακμές-κορυφές, οι μόνες των οποίων ο περιορισμός δεν ικανοποιείται στην ενδιάμεση ανάθεση ανήκουν σε 3-μονοπάτια στα οποία και οι δύο ακμές-κορυφές έχουν τιμή 1. Ο αριθμός αυτών των 3-μονοπατιών είναι τουλάχιστον  $t + 1 - s$  και το πολύ  $k + 1$ . Θεωρούμε τώρα το γράφημα  $H$  που αποτελείται από τις ακμές του  $G$  που αντιστοιχούν σε αυτά τα 3-μονοπάτια. Αφού το  $H$  έχει το πολύ  $k + 1$  ακμές και η περιφέρεια του  $G$  είναι τουλάχιστον  $k + 2$ , το  $H$  είναι ένα δάσος. Επιπλέον, οι κορυφές του γραφήματος  $H$  αντιστοιχούν στις  $t - s$  κορυφές που δεν είναι ακμές-κορυφές του  $G'$  των οποίων οι τιμές είναι 0 στην τελική ανάθεση. Το γράφημα  $H$  έχει τουλάχιστον  $t + 1 - s$  ακμές, που συνεπάγεται ότι περιέχει κύκλο. Επιπλέον, το  $H$  έχει το πολύ  $k + 1$  ακμές, επομένως η περιφέρειά του είναι το πολύ  $k + 1$ , κάτι που έρχεται σε αντίθεση με την αρχική υπόθεση.

Συμπερασματικά, καμία  $k$ -αλλαγή δε μπορεί να εφαρμοσθεί στην κακή ανάθεση που αναφέραμε προηγουμένως. Αποτελεί τοπικό βέλτιστο με κέρδος  $n(1 + d/2)$  ενώ το κέρδος της βέλτιστης ανάθεσης στο γράφημα  $G'$  είναι  $dn$ , που επιτυγχάνεται όταν ανατίθεται η τιμή 0 μόνο σε κορυφές που δεν είναι ακμές - κορυφές. Επομένως, ο λόγος προσέγγισης είναι το πολύ  $1/2 + 1/d \leq 1/2 + \epsilon$ .  $\square$



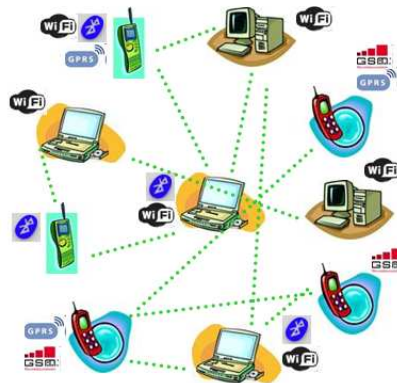
## Κεφάλαιο 4

# Ενεργειακά αποδοτική επικοινωνία σε ασύρματα δίκτυα με πολλαπλά μέσα ασύρματης διασύνδεσης

Στο Κεφάλαιο αυτό μελετάμε προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης. Σε τέτοια δίκτυα, δύο κόμβοι μπορούν να επικοινωνήσουν αν είναι αρκετά κοντά και διαθέτουν κάποιο κοινό μέσο ασύρματης διασύνδεσης. Η ενεργοποίηση ενός μέσου ασύρματης διασύνδεσης επιφέρει ένα κόστος που αντανακλά την ενέργεια που καταναλώνεται όταν κάποιος κόμβος χρησιμοποιεί το μέσο αυτό. Διακρίνουμε μεταξύ της συμμετρικής και της μη συμμετρικής περίπτωσης, με βάση τον το κόστος ενεργοποίησης για κάθε ασύρματο μέσο διασύνδεσης είναι το ίδιο για όλους τους κόμβους ή όχι. Για τη συμμετρική περίπτωση, παρουσιάζουμε έναν  $(3/2 + \epsilon)$ -προσεγγιστικό αλγόριθμο για το πρόβλημα πλήρους διασύνδεσης με ελάχιστο κόστος ενεργοποίησης, βελτιώνοντας ένα προηγούμενο φράγμα με τιμή 2. Για τη μη συμμετρική περίπτωση, αποδεικνύουμε ότι το πρόβλημα διασύνδεσης δεν είναι προσεγγίσιμο στα πλαίσια ενός παράγοντα υπολογαριθμικού ως προς το πλήθος των κόμβων και παρουσιάζουμε ένα λογαριθμικό προσεγγιστικό αλγόριθμο για μια γενικότερη περίπτωση που μοντελοποιεί την ομαδική επικοινωνία. Αυτά τα αποτελέσματα έχουν παρουσιαστεί στην εργασία [7].

## 4.1 Εισαγωγή

Κατά τα τελευταία χρόνια έχει δοθεί εξαιρετική προσοχή στα ασύρματα δίκτυα. Τα ασύρματα δίκτυα υποστηρίζουν ένα ευρύ φάσμα δημοφιλών εφαρμογών και συνήθως αποτελούν τμήματα μεγαλύτερων, καθολικών δικτύων, και του Διαδικτύου. Είναι γενικά ετερογενή με την έννοια ότι αποτελούνται από ασύρματες συσκευές με διαφορετικά χαρακτηριστικά όπως υπολογιστική ισχύ, κατανάλωση ενέργειας, μέσα ασύρματης διασύνδεσης, υποστηριζόμενα πρωτόκολλα επικοινωνίας, κτλ. Οι σύγχρονες ασύρματες συσκευές διαθέτουν πολλαπλά μέσα ασύρματης διασύνδεσης (όπως τα πιο ευρέως χρησιμοποιούμενα μέσα ασύρματης διασύνδεσης Bluetooth, WiFi και GPRS) και μπορούν να συνδέονται εναλλακτικά σε διαφορετικά δίκτυα επικοινωνίας με βάση απαιτήσεις συνεκτικότητας και περιορισμούς σχετικούς με ποιότητα υπηρεσιών (δείτε και Σχ. 4.1). Η επιλογή του καλύτερου μέσου ασύρματης διασύνδεσης ανά περίπτωση εξαρτάται από αρκετούς παράγοντες, όπως για παράδειγμα, διαθεσιμότητα ενός μέσου ασύρματης διασύνδεσης σε κάποια συγκεκριμένη συσκευή, περιορισμούς σχετικούς με παρεμβολές, το αναγκαίο εύρος ζώνης για επικοινωνία, την ενέργεια που καταναλώνεται από ένα ενεργοποιημένο μέσο ασύρματης διασύνδεσης και τη διάρκειά του, τα διαθέσιμα μέσα ασύρματης διασύνδεσης σε κάποια γειτονιά, ιδιότητες σχετικές με την τοπολογία του δικτύου, κτλ.



Σχήμα 4.1: Οι σύγχρονες ασύρματες συσκευές διαθέτουν πολλαπλά μέσα ασύρματης διασύνδεσης και μπορούν να συνδέονται σε διαφορετικά δίκτυα επικοινωνίας.

Μελετάμε προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρί-

ζουν πολλαπλά μέσα ασύρματης διασύνδεσης. Οι κόμβοι τέτοιων δικτύων είναι ασύρματες συσκευές που υποστηρίζουν κάποια μέσα ασύρματης διασύνδεσης. μεταξύ δύο τέτοιων κόμβων μπορεί να εγκαθιδρυθεί επικοινωνία αν (i) είναι αρκετά κοντά ο ένας στον άλλον και (ii) και οι δύο υποστηρίζουν κάποιο κοινό μέσο ασύρματης διασύνδεσης. Αν πληρούνται αυτές οι απαιτήσεις, τότε εγκαθιδρύεται επικοινωνία με κόστος ίσο με αυτό που απαιτείται για την ενεργοποίηση ενός συγκεκριμένου μέσου ασύρματης διασύνδεσης που υποστηρίζουν και οι δύο κόμβοι. Το κόστος ενεργοποίησης ενός μέσου ασύρματης διασύνδεσης αντανακλά την ενέργεια που καταναλώνεται όταν ένας κόμβος χρησιμοποιεί το μέσο αυτό. Στόχος μας είναι να ενεργοποιήσουμε μέσα ασύρματης διασύνδεσης στους κόμβους του δικτύου έτσι ώστε να διατηρείται κάποια ιδιότητα συνεκτικότητας και να ελαχιστοποιείται το συνολικό κόστος για τα ενεργοποιημένα μέσα ασύρματης διασύνδεσης. Ανάλογα με την απαιτούμενη ιδιότητα συνεκτικότητας, ανακύπτουν αρκετά προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης. Εστιάζουμε σε δύο τέτοια προβλήματα: ConMI και GroupMI. Στα πλαίσια του προβλήματος ConMI, απαιτούμε εγκαθίδρυση επικοινωνίας μεταξύ όλων των κόμβων του δικτύου. Στα πλαίσια του προβλήματος GroupMI, πρέπει να εγκαθιδρυθεί επικοινωνία μεταξύ ομάδων κόμβων (που δεν περιλαμβάνουν απαραίτητα όλους τους κόμβους του δικτύου). Το πρόβλημα ConMI αποτελεί ειδική περίπτωση του GroupMI. Διακρίνουμε δύο περιπτώσεις. Η γενικότερη είναι αυτή κατά την οποία το κόστος ενεργοποίησης για κάποιο μέσο ασύρματης διασύνδεσης δεν είναι το ίδιο για όλους τους κόμβους του δικτύου, την οποία καλούμε μη συμμετρική. Στη συμμετρική περίπτωση του προβλήματος, το κόστος ενεργοποίησης ενός συγκεκριμένου μέσου ασύρματης διασύνδεσης είναι το ίδιο σε όλους τους κόμβους του δικτύου.

**Σχετική δουλειά.** Τα ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης έχουν ιδιαίτερο ερευνητικό ενδιαφέρον αφού αποτελούν βασική επικοινωνιακή υποδομή και υποστηρίζουν ένα ευρύ φάσμα σημαντικών και δημοφιλών εφαρμογών. Με δεδομένο αυτό, πολλά βασικά προβλήματα που έχουν ήδη μελετηθεί για 'παραδοσιακά' ενσύρματα και ασύρματα δίκτυα έχουν επαναδιατυπωθεί [10], ειδικότερα εκείνα που σχετίζονται με ζητήματα συνεκτικότητας [20, 37] και δρομολόγησης [32]. Εντούτοις, οι απαιτήσεις σχετικά με αποδοτικότητα ως προς την ενέργεια αυξάνουν την πολυπλοκότητα των προβλημάτων αυτών και θέτουν νέες προκλήσεις.

Ένα συνδυαστικό πρόβλημα που εμπίπτει στη γενικότερη κατηγορία τω

προβλημάτων επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης μελετήθηκε στην εργασία [55]. Σε αυτή, δίνεται ένα γράφημα με επιθυμητές συνδέσεις μεταξύ κόμβων του δικτύου και στόχος είναι η ενεργοποίηση μέσω ασύρματης διασύνδεσης στους κόμβους του δικτύου με ελάχιστο συνολικό κόστος έτσι ώστε να εγκαθίστανται όλες οι ακμές του γραφήματος αυτού. Οι συγγραφείς μελετούν αρκετές παραλλαγές του προβλήματος αυτού που προκύπτουν με βάση την τοπολογία του γραφήματος εισόδου (δηλ., πλήρη γραφήματα, δένδρα, επίπεδα γραφήματα, γραφήματα φραγμένου βαθμού, γενικά γραφήματα) και με βάση το αν ο αριθμός των μέσων ασύρματης διασύνδεσης αποτελεί μέρος της εισόδου ή μια προκαθορισμένη σταθερά. Στην εργασία μελετώνται και μέσα ασύρματης διασύνδεσης με μοναδιαίο κόστος και γενικότερα συμμετρικά στιγμιότυπα.

Το πρόβλημα ConMI πρωτοεμφανίστηκε στην εργασία [57] όπου μελετήθηκαν συμμετρικά στιγμιότυπα του προβλήματος. Στην εργασία αποδεικνύεται ότι το πρόβλημα ConMI είναι δύσκολο να προσεγγιστεί (APX-hard) ακόμα και όταν το γράφημα που μοντελοποιεί το δίκτυο διαθέτει πολύ ιδιαίτερη δομή και ο αριθμός των διαθέσιμων μέσων ασύρματης διασύνδεσης είναι μικρός (δηλ., 2). Η θετική πλευρά είναι ότι στην εργασία [57] παρουσιάζεται ένας 2-προσεγγιστικός αλγόριθμος αξιοποιώντας τη σχέση του προβλήματος ConMI σε συμμετρικά στιγμιότυπα με το πρόβλημα εύρεσης ελάχιστου γεννητικού δένδρου σε κατάλληλα ορισμένο γράφημα με βάρη στις ακμές. Επιτυγχάνονται καλύτερα φράγματα προσέγγισης για ειδικές περιπτώσεις του προβλήματος ConMI όπως αυτή των μέσων ασύρματης διασύνδεσης μοναδιαίου κόστους .

**Αποτελέσματα.** Διακρίνουμε μεταξύ της συμμετρικής και της μη συμμετρικής περίπτωσης, με βάση το αν όλοι οι κόμβοι του δικτύου έχουν το ίδιο κόστος ενεργοποίησης για κάποιο μέσο ασύρματης διασύνδεσης ή όχι. Για τη συμμετρική περίπτωση, παρουσιάζουμε έναν  $(3/2 + \epsilon)$ -προσεγγιστικό αλγόριθμο για το πρόβλημα ConMI, βελτιώνοντας το καλύτερο γνωστό φράγμα με τιμή 2 από την εργασία [57]. Η βασική ιδέα του αλγορίθμου είναι να χρησιμοποιεί ένα σχεδόν ελάχιστο γεννητικό δένδρο (minimum spanning tree-MST) σε ένα κατάλληλα ορισμένο υπεργράφημα και να το μετατρέπει σε μια αποδοτική λύση για το πρόβλημα ConMI. Μελετάμε επίσης το πρόβλημα GroupMI για συμμετρικά στιγμιότυπα για το οποίο παρουσιάζουμε έναν 4-προσεγγιστικό αλγόριθμο. Τώρα, μετατρέπουμε στιγμιότυπα του προβλήματος σε αντίστοιχα του προβλήματος εύρεσης δασών Steiner με παρόμοιο

τρόπο με αυτόν που στην εργασία [57] μετατρέπεται το πρόβλημα ConMI στο πρόβλημα MST. Για τη μη συμμετρική περίπτωση, αποδεικνύουμε ότι το πρόβλημα συνεκτικότητας δεν είναι προσεγγίσιμο στα πλαίσια ενός παράγοντα υπολογαριθμικού ως προς το πλήθος των κόβων μέσω αναγωγής από το πρόβλημα Κάλυψης Συνόλου (Set Cover), και παρουσιάζουμε έναν λογαριθμικό προσεγγιστικό αλγόριθμο για το γενικότερο πρόβλημα GroupMI. Μετατρέπουμε στιγμιότυπα του προβλήματος σε αντίστοιχα του προβλήματος εύρεσης δάσους Steiner σε γραφήματα με βάρη στις κορυφές και αξιοποιούμε προσεγγιστικούς αλγόριθμους των Guha και Khuller [41] (επίσης [2]) για εύρεση δάσους Steiner με βάρη στις κορυφές. Αξίζει να σημειωθεί ότι τεχνικές για εύρεση δάσους Steiner με βάρη στις κορυφές έχουν χρησιμοποιηθεί είτε έμμεσα [13] είτε ρητά [17] για προβλήματα επικοινωνίας με ελάχιστη κατανάλωση ενέργειας σε ασύρματα αδόμητα δίκτυα (ασύρματα δίκτυα τύπου ad hoc). Με βάση όσα γνωρίζουμε, τόσο το πρόβλημα GroupMI όσο και μη συμμετρικά του στιγμιότυπα σε ασύρματα δίκτυα με πολλαπλά μέσα ασύρματης διασύνδεσης δεν έχουν μελετηθεί στο παρελθόν.

Το υπόλοιπο Κεφάλαιο ακολουθεί την εξής δομή. Στην Ενότητα 4.2, παρουσιάζουμε αρχικούς τεχνικούς ορισμούς και συμβολισμούς που χρησιμοποιούμε. Στην Ενότητα 4.3 παρουσιάζουμε το άνω φράγμα για συμμετρικά στιγμιότυπα του προβλήματος ConMI, ενώ στην Ενότητα 4.4 παραθέτουμε τον αλγόριθμο για το πρόβλημα GroupMI. Τα αποτελέσματα για μη συμμετρικά στιγμιότυπα του προβλήματος GroupMI παρουσιάζονται στην Ενότητα 4.5.

## 4.2 Ορισμοί και συμβολισμοί

Το ασύρματο δίκτυο μοντελοποιείται μέσω ενός γραφήματος  $G$  του οποίου οι κορυφές αντιστοιχούν σε κόμβους του δικτύου και οι ακμές αναπαριστούν ενδεχόμενες απευθείας συνδέσεις μεταξύ ζευγών κόμβων του δικτύου. Συμβολίζουμε με  $I$  το σύνολο των διαθέσιμων μέσων ασύρματης διασύνδεσης. Κάθε κορυφή  $u$  υποστηρίζει ένα σύνολο  $I_u \subseteq I$  μέσων ασύρματης διασύνδεσης. Δύο κορυφές  $u$  και  $v$  μπορούν να επικοινωνήσουν όταν έχουν ενεργοποιημένο το ίδιο μέσο ασύρματης διασύνδεσης αρκεί να υπάρχει μεταξύ τους ακμή στο γράφημα  $G$ . Δεδομένου ενός συνόλου ενεργοποιημένων μέσων ασύρματης διασύνδεσης  $S_u \subseteq I_u$  σε κάθε κορυφή  $u$ , ορίζουμε το γράφημα επικοινωνίας  $G_S$  που έχει το ίδιο σύνολο κορυφών με το γράφημα  $G$  και μια ακμή  $e = (u, v)$  του  $G$  ανήκει στο  $G_S$  αν  $S_u \cap S_v \neq \emptyset$ . Η ενεργοποίηση του μέσου ασύρματης διασύνδεσης  $g$  στην κορυφή  $u$  έχει ένα μη αρνητικό κόστος

$c_{u,g}$ . Στόχος μας είναι να ενεργοποιήσουμε μέσα ασύρματης διασύνδεσης στις κορυφές του  $G$  έτσι ώστε το αντίστοιχο γράφημα επικοινωνίας να διατηρεί κάποια ιδιότητα συνεκτικότητας και να ελαχιστοποιείται το συνολικό κόστος των ενεργοποιημένων μέσων ασύρματης διασύνδεσης. Ανάλογα με την επιθυμητή ιδιότητα συνεκτικότητας, ανακύπτουν αρκετά προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζου πολλαπλά μέσα ασύρματης διασύνδεσης. Μελετούμε δύο τέτοια προβλήματα: το ConMI και το GroupMI. Στα πλαίσια του προβλήματος ConMI, απαιτούμε το γράφημα επικοινωνίας να είναι συνεκτικό και να περιέχει όλες τις κορυφές του  $G$ . Στα πλαίσια του προβλήματος GroupMI, έχουμε επιπλέον σα δεδομένο ένα σύνολο τερματικών κορυφών  $D \subseteq V$  διαμερισμένο σε  $p$  ξένα μεταξύ τους υποσύνολα  $D_1, D_2, \dots, D_p$ . Για  $i = 1, \dots, p$ , στο γράφημα επικοινωνίας πρέπει να συνδέονται οι τερματικές κορυφές του συνόλου  $D_i$ . Προφανώς, το πρόβλημα ConMI αποτελεί ειδική περίπτωση του προβλήματος GroupMI. Διακρίνουμε δύο περιπτώσεις. Η γενικότερη που περιγράφηκε παραπάνω είναι η μη συμμετρική. Στη συμμετρική περίπτωση του προβλήματος, το κόστος ενεργοποίησης του μέσου ασύρματης διασύνδεσης  $g$  σε κάθε κορυφή είναι το ίδιο και ίσο με  $c_g$ .

Στη συνέχεια, αναφερόμαστε συχνά σε γνωστά προβλήματα συνδυαστικής βελτιστοποίησης όπως το πρόβλημα εύρεσης δάσους Steiner και το πρόβλημα εύρεσης δάσους Steiner με βάρη στις κορυφές. Και στα δύο προβλήματα, η είσοδος αποτελείται από ένα γράφημα  $G = (V, E)$  και ένα σύνολο τερματικών  $D \subseteq V$  διαμερισμένο σε  $p$  ξένα μεταξύ τους υποσύνολα όπως στο πρόβλημα GroupMI, και στόχος είναι ο υπολογισμός ενός δάσους ελάχιστου κόστους (βάρους) έτσι ώστε τα τερματικά στο υποσύνολο  $D_i$  να ανήκουν στο ίδιο δένδρο του δάσους. Στο πρόβλημα εύρεσης δάσους Steiner, κάθε ακμή  $e$  του  $G$  έχει ένα μη αρνητικό βάρος  $w_e$ . Στο πρόβλημα εύρεσης δάσους Steiner με βάρη στις κορυφές, οι ακμές δεν έχουν βάρη και κάθε κορυφή  $u$  έχει κάποιο βάρος  $w_u$ .

### 4.3 Ένας αλγόριθμος για τη συμμετρική περίπτωση

Παρουσιάζουμε έναν  $(3/2 + \epsilon)$ -προσεγγιστικό αλγόριθμο για τη συμμετρική περίπτωση του προβλήματος ConMI βελτιώνοντας το προηγούμενο γνωστό άνω φράγμα με τιμή 2.

Ο αλγόριθμός μας λειτουργεί ως εξής. Θεωρούμε ένα στιγμιότυπο εισό-



δου  $J$ , ένα σύνολο διαθέσιμων μέσων ασύρματης διασύνδεσης  $I$  και σύνολα  $I_u$  μέσων ασύρματης διασύνδεσης που υποστηρίζονται από κάθε κορυφή  $u$ . Αρχικά, μετατρέπουμε το γράφημα  $G$  σε στιγμιότυπο του προβλήματος υπολογισμού ελάχιστου γεννητικού δένδρου (MST) σε ένα κατάλληλα ορισμένο υπεργράφημα  $H$ . Στη συνέχεια, λύνουμε σχεδόν βέλτιστα το πρόβλημα MST στο γράφημα  $H$  χρησιμοποιώντας ένα σχήμα προσέγγισης πολυωνυμικού χρόνου των Prömel και Steger [68]. Χρησιμοποιούμε το δένδρο που προκύπτει για να καθορίσουμε ποια μέσα ασύρματης διασύνδεσης θα ενεργοποιήσουμε σε κάθε κορυφή του  $G$  έτσι ώστε το αντίστοιχο γράφημα επικοινωνίας να είναι ένα συνεκτικό γεννητικό υπογράφημα του  $G$ . Αυτό είναι και ό,τι επιστρέφει ο αλγόριθμός μας. Αποδεικνύουμε ότι ο αλγόριθμος επιτυγχάνει εγγύηση προσέγγισης  $3/2 + \epsilon$ , όπου  $\epsilon$  είναι η εγγύηση απόδοσης του MST αλγορίθμου στο υπεργράφημα  $H$ .

Το υπεργράφημα  $H = (V, F)$  έχει το ίδιο σύνολο κορυφών με το γράφημα  $G$ . Ορίζουμε το σύνολο των ακμών  $F$  του  $H$  ως εξής. Θεωρούμε όλες τις τριάδες κορυφών  $v_i, v_j, v_k$  τέτοιες ώστε  $I_{v_i} \cap I_{v_j} \cap I_{v_k} \neq \emptyset$  που συνδέονται με τουλάχιστον δύο ακμές μεταξύ τους στο  $G$ . Εισάγουμε την τριάδα  $(v_i, v_j, v_k)$  σε μια υπερακμή  $f$  του  $F$ . Συμβολίζουμε με  $s(f)$  το μέσο ασύρματης διασύνδεσης με το ελάχιστο κόστος στο σύνολο  $I_{v_i} \cap I_{v_j} \cap I_{v_k}$ . Καλούμε  $s(f)$  το μέσο ασύρματης διασύνδεσης που σχετίζεται με την υπερακμή  $f$ . Αναθέτουμε στην  $f$  βάρος  $3c_{s(f)}$ . Αυτό αντιστοιχεί στο γεγονός ότι ενεργοποιώντας το μέσο ασύρματης διασύνδεσης  $s(f)$  στις κορυφές  $v_i, v_j, v_k$ , οι ακμές που τις συνδέουν περιέχονται στο γράφημα επικοινωνίας με κόστος  $3c_{s(f)}$ .

Θεωρούμε όλα τα ζεύγη κορυφών  $v_i, v_j$  τέτοια ώστε  $I_{v_i} \cap I_{v_j} \neq \emptyset$  που συνδέονται με μια ακμή στο  $G$ . Εισάγουμε το ζεύγος  $(v_i, v_j)$  σε μια υπερακμή  $f$  του  $F$ . Πάλι, συμβολίζουμε με  $s(f)$  το μέσο ασύρματης διασύνδεσης με ελάχιστο κόστος στο  $I_{v_i} \cap I_{v_j}$ . Αναθέτουμε στην  $f$  βάρος  $2c_{s(f)}$ . Αυτό αντιστοιχεί στο γεγονός ότι ενεργοποιώντας το μέσο  $s(f)$  στις κορυφές  $v_i, v_j$ , η ακμή που τις συνδέει περιέχεται στο αντίστοιχο γράφημα επικοινωνίας με κόστος  $2c_{s(f)}$ .

Αφού οι ακμές του υπεργραφήματος  $H$  αποτελούνται από το πολύ 3 κορυφές, χρησιμοποιούμε το σχήμα προσέγγισης πολυωνυμικού χρόνου της εργασίας [68] για να πάρουμε ένα γεννητικό δένδρο  $T$  του  $H$ . Για κάθε ακμή  $f$  του  $T$ , ενεργοποιούμε το μέσο ασύρματης διασύνδεσης  $s(f)$  στις κορυφές του  $G$  που ανήκουν στην  $f$ . Έτσι, στο αντίστοιχο γράφημα επικοινωνίας οι κορυφές που ανήκουν στην ίδια υπερακμή του  $T$  συνδέονται και επειδή το  $T$  είναι συνεκτικό και χρησιμοποιεί όλες τις κορυφές του  $V$ , το συνολικό γράφημα επικοινωνίας είναι, επίσης, ένα γεννητικό υπογράφημα του  $G$ .

Συμβολίζουμε με  $cost(J)$  το συνολικό κόστος της λύσης που επιτυγχάνεται

από τον αλγόριθμό μας, με  $opt(J)$  το κόστος της βέλτιστης λύσης, με  $mst(H)$  το κόστος του ελάχιστου γεννητικού δένδρου του  $H$  και με  $st(H)$  το κόστος του γεννητικού δένδρου  $T$ . Αποδεικνύουμε τα ακόλουθα δύο λήμματα.

**Λήμμα 14.**  $cost(J) \leq st(H)$ .

*Απόδειξη.* Επειδή το σύνολο  $S_u$  των μέσων ασύρματης διασύνδεσης που είναι ενεργοποιημένα σε κάθε κορυφή  $u$  αποτελείται από μέσα συσχετισμένα με τις υπερακμές του  $T$  που περιέχουν την  $u$ , ισχύει ότι:

$$cost(J) = \sum_{u \in V} \sum_{g \in S_u} c_g \leq \sum_{u \in V} \sum_{f \in T: u \in f} c_{s(f)} = \sum_{f \in T} w(f) = st(H).$$

□

**Λήμμα 15.**  $mst(H) \leq \frac{3}{2}opt(J)$ .

*Απόδειξη.* Θεωρούμε μια βέλτιστη λύση για ένα στιγμιότυπο  $J$  του προβλήματος ConMI που αποτελείται από σύνολα μέσων ασύρματης διασύνδεσης  $S_u$  που είναι ενεργοποιημένα σε κάθε κορυφή  $u$  του  $G$ . Συμβολίζουμε με  $S$  το σύνολο όλων των ενεργοποιημένων μέσων ασύρματης διασύνδεσης και με  $G_S$  το αντίστοιχο γράφημα επικοινωνίας. Αποσυνθέτουμε το γράφημα  $G_S$  σε διαφορετικά υπογράφημα. Υπάρχει ένα τέτοιο υπογράφημα για κάθε μέσο ασύρματης διασύνδεσης του  $S$ . Συμβολίζουμε με  $G_g$  το υπογράφημα του  $G$  που αποτελείται από το σύνολο κορυφών  $V_g$  στις οποίες είναι ενεργοποιημένο το μέσο ασύρματης διασύνδεσης  $g$  στη βέλτιστη λύση και από το σύνολο ακμών  $E_g$  που συνδέουν κορυφές του  $V_g$  στο  $G$ . Προφανώς,  $opt(J) = \sum_{g \in S} c_g |V_g|$ . Για κάθε  $g \in S$ , υπολογίζουμε ένα ελάχιστο γεννητικό δένδρο σε κάθε συνεκτική συνιστώσα του  $G_g$ . Τα ελάχιστα γεννητικά δένδρα στις συνεκτικές συνιστώσες του  $G_g$  σχηματίζουν ένα δάσος  $T_g$ .

Αποσυνθέτουμε τις ακμές του  $T_g$  σε ιδιαίτερες υποδομές που καλούμε *πηρούνια*: ένα πηρούνι είναι είτε ένα σύνολο δύο ακμών που πρόσκεινται στην ίδια κορυφή είτε μια μόνο ακμή. Σε κάθε συνεκτική συνιστώσα του  $T_g$  με  $m$  κορυφές, η διαδικασία που αποσυνθέτει τις ακμές της σε πηρούνια είναι η εξής. Αν υπάρχουν δύο φύλλα  $u$  και  $v$  με κοινό γονέα, συμπεριλαμβάνουμε τις ακμές που πρόσκεινται στις  $u$  και  $v$  σε ένα πηρούνι. Αφαιρούμε τις  $u$ ,  $v$ , και τις προσκείμενες ακμές τους από το δένδρο. Αν δεν υπάρχουν φύλλα με κοινό γονέα, τότε κάποιο φύλλο  $u$  έχει σα γονέα μια κορυφή  $v$  με βαθμό 2 ή έχει απομείνει μία μόνο ακμή μεταξύ δύο κορυφών  $u$  και  $v$ . Στην πρώτη περίπτωση, συμπεριλαμβάνουμε τις ακμές που πρόσκεινται στις κορυφές  $u$

και  $v$  σε ένα πηρούνι και αφαιρούμε τις  $u, v$ , και τις προσκείμενες ακμές τους από το δένδρο. Στη δεύτερη περίπτωση, απλά συμπεριλαμβάνουμε την ακμή μεταξύ των  $u$  και  $v$  σε ένα πηρούνι και την αφαιρούμε από το δένδρο. Επαναλαμβάνουμε την παραπάνω διαδικασία έως ότου όλες οι ακμές του δένδρου να έχουν συμπεριληφθεί σε πηρούνια. Σε κάθε βήμα (ενδεχομένως εκτός του τελευταίου), 2 από τις το πολύ  $m - 1$  ακμές του δένδρου συμπεριλαμβάνονται σε ένα πηρούνι. Επομένως, ο αριθμός των πηρουνιών είναι το πολύ  $m/2$ . Επαναλαμβάνοντας αυτή την αποσύνθεση για κάθε συνεκτική συνιστώσα του  $T_g$ , λαμβάνουμε μια αποσύνθεση των ακμών του  $T_g$  σε το πολύ  $|V_g|/2$  πηρούνια. Τα άκρα κάθε πηρουνιού του  $T_g$  αντιστοιχούν σε μια υπερακμή στο γράφημα  $H$  με βάρος το πολύ  $3c_g$ . Η ένωση όλων αυτών των υπερακμών είναι ένα συνεκτικό γεννητικό υπογράφημα του  $H$  (αφού η ένωση των  $T_g$  δίνει το  $G_S$ ). Το κόστος του ελάχιστου γεννητικού δένδρου του  $H$  φράσσεται άνω από το συνολικό κόστος των υπερακμών σε αυτό το γεννητικό υπογράφημα, δηλ.,

$$mst(H) \leq \sum_{g \in S} 3c_g \frac{|V_g|}{2} = \frac{3}{2} opt(J).$$

□

Από τα Λήμματα 14 και 15 και αφού  $st(H) \leq (1 + \epsilon)mst(H)$ , καταλήγουμε ότι:

**Θεώρημα 16.** *Για κάθε σταθερά  $\epsilon > 0$ , υπάρχει ένας  $(3/2 + \epsilon)$ -προσεγγιστικός αλγόριθμος πολυωνυμικού χρόνου για το πρόβλημα ConMI.*

## 4.4 Ομαδική επικοινωνία στη συμμετρική περίπτωση

Στην ενότητα αυτή, παρουσιάζουμε έναν αλγόριθμο για συμμετρικά στιγμιότυπα του προβλήματος GroupMI που επιτυγχάνει σταθερό λόγο προσέγγισης. Η βασική ιδέα του αλγορίθμου είναι ίδια με τον αλγόριθμο του Kosowski και άλλων [57] για το πρόβλημα ConMI αλλά αντί να χρησιμοποιούμε ένα αλγόριθμο πολυωνυμικού χρόνου για το πρόβλημα MST, χρησιμοποιούμε τον 2-προσεγγιστικό αλγόριθμο των Goemans και Williamson [39] για το πρόβλημα εύρεσης δάσους Steiner.

Θεωρούμε ένα στιγμιότυπο  $J$  του προβλήματος GroupMI με ένα γράφημα  $G = (V, E)$  με  $n$  κορυφές, ένα σύνολο τερματικών κορυφών  $D \subseteq V$  διαμερισμένο σε  $p$  ξένα μεταξύ τους υποσύνολα  $D_1, \dots, D_p$  και σύνολα  $I_u$  μέσων

ασύρματης διασύνδεσης που υποστηρίζονται από κάθε κορυφή  $u \in V$ . Κατασκευάζουμε ένα στιγμιότυπο  $J_{SF}$  του προβλήματος εύρεσης δάσους Steiner που αποτελείται από ένα γράφημα  $H = (V, A)$  και το σύνολο των τερματικών  $D$  διαμερισμένο στα ίδια  $p$  ζένα μεταξύ τους υποσύνολα τερματικών. Το σύνολο ακμών  $A$  περιέχει όλες τις ακμές  $(u, v)$  του  $E$  έτσι ώστε  $I_u \cap I_v \neq \emptyset$ . Θεωρούμε μια ακμή  $e = (u, v)$  του  $A$  και έστω  $s(e)$  το μέσο ασύρματης διασύνδεσης ελάχιστου κόστους στο  $I_u \cap I_v$ . Τότε, το βάρος  $w_e$  της  $e$  ισούται με  $2c_{s(e)}$ .

Χρησιμοποιούμε τον αλγόριθμο της εργασίας [39] για να λύσουμε το πρόβλημα εύρεσης δάσους Steiner για το στιγμιότυπο  $J_{SF}$  και λαμβάνουμε ένα δάσος  $F$  που διατηρεί συνεκτικότητα μεταξύ των κορυφών κάθε τερματικού συνόλου  $D_i$ . Επιτυγχάνουμε τη λύση στο αρχικό στιγμιότυπο  $J$  ως εξής. Για κάθε ακμή  $e$  στο  $F$ , ενεργοποιούμε το μέσο ασύρματης διασύνδεσης  $s(e)$  στα άκρα της  $e$ . Προφανώς, με τον τρόπο αυτό οι ακμές του γραφήματος  $G$  που αντιστοιχούν στην  $F$  περιέχονται στο αντίστοιχο γράφημα επικοινωνίας και ικανοποιούνται οι απαιτούμενες απαιτήσεις συνεκτικότητας για το στιγμιότυπο  $J$ .

Το άνω φράγμα για το λόγο προσέγγισης του αλγορίθμου δίνεται στο ακόλουθο θεώρημα. Η απόδειξη προκύπτει επεκτείνοντας τα επιχειρήματα που χρησιμοποιούνται στην εργασία [57].

**Θεώρημα 17.** *Υπάρχει ένας 4-προσεγγιστικός αλγόριθμος για συμμετρικά στιγμιότυπα του προβλήματος GroupMI.*

## 4.5 Ομαδική επικοινωνία στη μη συμμετρική περίπτωση

Στην ενότητα αυτή, μελετάμε το πρόβλημα GroupMI για μη συμμετρικά στιγμιότυπα. Στην περίπτωση αυτή, ακόμα και για το απλούστερο πρόβλημα ConMI δεν υπάρχουν προσεγγιστικοί αλγόριθμοι σταθερής προσέγγισης όπως υποδεικνύει η ακόλουθη δήλωση.

**Θεώρημα 18.** *Η μη συμμετρική περίπτωση του προβλήματος ConMI σε δίκτυα με  $n$  κόμβους είναι δύσκολο να προσεγγιστεί στα πλαίσια ενός παράγοντα  $o(\ln n)$ .*

*Απόδειξη.* Χρησιμοποιούμε μια απλή αναγωγή από το πρόβλημα Κάλυψης Συνόλου. Θεωρούμε ένα στιγμιότυπο του προβλήματος Κάλυψης Συνόλου με

σύνολο βάσης  $U$  με  $m$  στοιχεία και μια συλλογή  $\mathcal{T}$  από υποσύνολα του  $U$ . Το μέγεθος της συλλογής  $\mathcal{T}$  είναι πολυωνυμικό ως προς  $m$ . Κατασκευάζουμε ένα στιγμιότυπο του προβλήματος ConMI ως εξής. Το σύνολο των μέσων ασύρματης διασύνδεσης  $I$  περιέχει δύο τέτοια μέσα 0 και 1. Το γράφημα  $G$  έχει μια κορυφή ρίζα  $r$ , κορυφές  $u_1, \dots, u_{|\mathcal{T}|}$  που αντιστοιχούν στα σύνολα του  $\mathcal{T}$ , και κορυφές  $v_1, \dots, v_m$  που αντιστοιχούν στα στοιχεία του  $U$ . Η κορυφή  $r$  υποστηρίζει μόνο το μέσο ασύρματης διασύνδεσης 0 (δηλ.,  $I_r = \{0\}$ ) με κόστος ενεργοποίησης 0. Οι κορυφές  $u_1, \dots, u_{|\mathcal{T}|}$  υποστηρίζουν το μέσο 0 και 1 (δηλ.,  $I_{u_i} = \{0, 1\}$ ) με κόστος ενεργοποίησης 0 για μέσο ασύρματης διασύνδεσης 0 και 1 για μέσο ασύρματης διασύνδεσης 1. Για  $i = 1, \dots, m$ , η κορυφή  $v_i$  υποστηρίζει μόνο το μέσο ασύρματης διασύνδεσης 1 ( $I_{v_i} = \{1\}$ ) με κόστος ενεργοποίησης 0. Για κάθε σύνολο  $T_i$  του  $\mathcal{T}$ , η κορυφή  $u_i$  συνδέεται με ακμή με κάθε κορυφή  $v_j$  έτσι ώστε το στοιχείο  $j$  να ανήκει στο σύνολο  $T_i$ . Η κορυφή ρίζα  $r$  έχει ακμές προς κάθε κορυφή  $u_i$ , για  $i = 1, \dots, |\mathcal{T}|$ .

Μπορούμε εύκολα να δείξουμε ότι κάθε κάλυψη του  $U$  με  $C$  σύνολα από το  $\mathcal{T}$  δίνει μια λύση για το πρόβλημα ConMI με κόστος το πολύ  $C$  και αντίστροφα. Πράγματι, θεωρούμε μια λύση για ένα στιγμιότυπο του προβλήματος Κάλυψης Συνόλου που αποτελείται από ένα υποσύνολο  $\mathcal{T}'$  του  $\mathcal{T}$ . Ενεργοποιώντας το μέσο ασύρματης διασύνδεσης 0 στις κορυφές  $r, u_1, \dots, u_{\mathcal{T}}$  και το μέσο ασύρματης διασύνδεσης 1 στις κορυφές  $v_1, \dots, v_m$  και τις κορυφές  $u_i$ , έτσι ώστε  $i \in \mathcal{T}'$ , λαμβάνουμε μια λύση στο στιγμιότυπο του προβλήματος ConMI με κόστος  $|\mathcal{T}'|$ . Επίσης, δεδομένης μιας λύσης για ένα στιγμιότυπο του προβλήματος ConMI, λαμβάνουμε μια κάλυψη του  $U$  ίδιου κόστους επιλέγοντας τα σύνολα του  $\mathcal{T}$  που αντιστοιχούν στις κορυφές  $u_i$  στις οποίες είναι ενεργοποιημένο το μέσο ασύρματης διασύνδεσης 1. Χρησιμοποιώντας γνωστά αποτελέσματα μη προσεγγισιμότητας για το πρόβλημα Κάλυψης Συνόλου [38, 72], επιτυγχάνουμε ένα φράγμα μη προσεγγισιμότητας με τιμή  $\tau \ln m$ . Αφού το πλήθος των κορυφών  $n$  στο στιγμιότυπο του προβλήματος ConMI είναι πολυωνυμικό ως προς  $m$ , καταλήγουμε στο ζητούμενο.  $\square$

Στη συνέχεια, παρουσιάζουμε έναν  $O(\ln n)$ -προσεγγιστικό αλγόριθμο για τη μη συμμετρική περίπτωση του προβλήματος GroupMI ανάγοντας το πρόβλημα σε στιγμιότυπα του προβλήματος εύρεσης δάσους Steiner με βάρη στις κορυφές. Η αναγωγή μοιάζει με αντίστοιχες για προβλήματα επικοινωνίας με ελάχιστη ενέργεια σε ασύρματα αδόμητα δίκτυα (ασύρματα δίκτυα τύπου ad hoc) [17].

Θεωρούμε ένα στιγμιότυπο  $J$  του προβλήματος GroupMI με ένα γράφημα  $G = (V, E)$  με  $n$  κορυφές, ένα σύνολο τερματικών κορυφών  $D \subseteq V$  διαμε-

ρισμένο σε  $p$  ξένα μεταξύ τους υποσύνολα  $D_1, \dots, D_p$  και σύνολα  $I_u$  μέσω ασύρματης διασύνδεσης που υποστηρίζονται από κάθε κορυφή  $u \in V$ . Κατασκευάζουμε ένα στιγμιότυπο του προβλήματος εύρεσης δάσους Steiner με βάρη στις κορυφές που αποτελείται από ένα γράφημα  $H = (U, A)$  και ένα σύνολο τερματικών  $D' \subseteq U$  διαμερισμένο σε  $p$  ξένα μεταξύ τους υποσύνολα  $D'_1, \dots, D'_p$ . Το γράφημα  $H$  ορίζεται ως εξής. Το σύνολο κορυφών  $U$  αποτελείται από  $n$  ξένα μεταξύ τους σύνολα κορυφών που καλούνται υπερκορυφές. Κάθε υπερκορυφή αντιστοιχεί σε μια κορυφή του  $V$ . Η υπερκορυφή  $Z_u$  που αντιστοιχεί στην κορυφή  $u \in V$  περιέχει τις ακόλουθες  $|I_u| + 1$  κορυφές: μια κεντρική κορυφή  $Z_{u,0}$  και  $|I_u|$  κορυφές-γέφυρες  $Z_{u,g}$  για κάθε μέσο ασύρματης διασύνδεσης  $g \in I_u$ . Για κάθε ζεύγος κορυφών  $u, v \in V$  και κάθε μέσο ασύρματης διασύνδεσης  $g \in I_u \cap I_v$ , το σύνολο  $A$  των ακμών περιέχει μία ακμή μεταξύ των κορυφών-γεφυρών  $Z_{u,g}$  και  $Z_{v,g}$ . Επίσης, για κάθε κορυφή  $u \in V$ , το  $A$  περιέχει μια ακμή μεταξύ της κεντρικής κορυφής  $Z_{u,0}$  και κάθε κορυφής-γέφυρας  $Z_{u,g}$ , για  $g \in I_u$ . Κάθε κεντρική κορυφή έχει βάρος 0. Μια κορυφή-γέφυρα  $Z_{u,g}$  που αντιστοιχεί στην κορυφή  $u \in V$  και στο μέσο ασύρματης διασύνδεσης  $g \in I_u$  έχει βάρος ίσο με το κόστος ενεργοποίησης  $c_{u,g}$  του μέσου ασύρματης διασύνδεσης  $g$  στην κορυφή  $u$ . Το σύνολο των τερματικών  $D'$  αποτελείται από όλες τις κεντρικές κορυφές. Για  $i = 1, \dots, p$ , το σύνολο  $D'_i$  στη διαμέριση του  $D'$  περιέχει τις κεντρικές κορυφές  $Z_{u,0}$  για κάθε κορυφή  $u \in D_i$ .

Συμβολίζουμε με  $J_{NWSF}$  το στιγμιότυπο του προβλήματος εύρεσης δάσους Steiner με βάρη στις κορυφές που προκύπτει. Χρησιμοποιούμε ένα γνωστό αλγόριθμο για να λύσουμε το πρόβλημα εύρεσης δάσους Steiner με βάρη στις κορυφές για το στιγμιότυπο  $J_{NWSF}$  και λαμβάνουμε ένα δάσος  $F$  που είναι υπογράφημα του  $H$  χωρίς απομονωμένες κορυφές και διατηρεί συνεκτικότητα μεταξύ των κορυφών κάθε τερματικού συνόλου  $D_i$ . Λαμβάνουμε τη λύση  $S$  για το αρχικό στιγμιότυπο  $J$  ως εξής. Για κάθε μέσο ασύρματης διασύνδεσης  $g$  στο  $I_u$ , περιλαμβάνουμε το  $g$  στο  $S_u$  αν και μόνον αν η  $Z_{u,g}$  είναι κορυφή του  $F$ .

Το ακόλουθο λήμμα περιλαμβάνει τη βασική ιδιότητα της αναγωγής.

**Λήμμα 19.** *Αν η  $F$  είναι μια  $\rho$ -προσεγγιστική λύση για το  $J_{NWSF}$ , τότε η  $S$  είναι μια  $\rho$ -προσεγγιστική λύση για το  $J$ .*

*Απόδειξη.* Έστω  $u, v$  κορυφές που ανήκουν στο ίδιο τερματικό σύνολο  $D_i$ . Τότε, οι κεντρικές κορυφές  $Z_{u,0}$  και  $Z_{v,0}$  ανήκουν στο σύνολο  $D'_i$  και υπάρχει μονοπάτι  $q$  από τη  $Z_{u,0}$  στη  $Z_{v,0}$  στο γράφημα  $F$ . Οι ακμές του  $q$  είναι είτε ακμές που συνδέουν μια κεντρική κορυφή με μια κορυφή-γέφυρα στην ίδια

υπερκορυφή ή κορυφές-γέφυρες μεταξύ διαφορετικών υπερκορυφών. Θεωρούμε τις ακμές του  $q$  που συνδέουν κορυφές-γέφυρες διαφορετικών υπερκορυφών με τη σειρά που τις επισκεπτόμαστε ακολουθώντας το μονοπάτι  $q$  από τη  $Z_{u,0}$  στην  $Z_{v,0}$ . Για κάθε τέτοια ακμή  $(Z_{x,g}, Z_{y,g})$  η ακμή  $(x, y)$  ανήκει στο γράφημα  $G$  και όλες μαζί ορίζουν ένα μονοπάτι  $q'$  από τη  $u$  στη  $v$  στο γράφημα  $G$ . Επειδή οι  $Z_{x,g}$  και  $Z_{y,g}$  ανήκουν στην  $F$ , στις κορυφές  $x$  και  $y$  είναι ενεργοποιημένο το μέσο ασύρματης διασύνδεσης  $g$  και επομένως οι ακμές του  $q'$  ανήκουν στο γράφημα επικοινωνίας που δημιουργείται.

Στη συνέχεια δείχνουμε ότι το συνολικό κόστος ενεργοποίησης  $cost(J)$  της λύσης μας ισούται με το κόστος  $cost(J_{NWSF})$  της  $F$  και ότι το βέλτιστο κόστος ενεργοποίησης  $opt(J)$  φράσσεται κάτω από το κόστος  $opt(J_{NWSF})$  της βέλτιστης λύσης για το  $J_{NWSF}$ . Με τον τρόπο αυτό, έχουμε ότι

$$cost(J) = cost(J_{NWSF}) \leq \rho \cdot opt(J_{NWSF}) \leq \rho \cdot opt(J).$$

Πράγματι, το μέσο ασύρματης διασύνδεσης  $g$  ενεργοποιείται στην κορυφή  $u$  μόνο αν η κορυφή-γέφυρα  $Z_{u,g}$  ανήκει στην  $F$ . Επειδή,  $w(Z_{u,g}) = c_{u,g}$ , έχουμε ότι το συνολικό κόστος ενεργοποίησης της λύσης μας για το στιγμιότυπο  $J$  ισούται με το κόστος της λύσης  $F$ .

Θεωρούμε τώρα μια βέλτιστη λύση για το  $J$  που αποτελείται από σύνολα  $S_u$  ενεργοποιημένων μέσων ασύρματης διασύνδεσης σε κάθε κορυφή  $u$  του γραφήματος  $G$ . κατασκευάζουμε ένα υπογράφημα  $F'$  του  $H$  ως εξής. Για κάθε ακμή  $(u, v)$  στο γράφημα επικοινωνίας  $G_S$ , και για κάθε μέσο ασύρματης διασύνδεσης  $g$  που ανήκει στο  $S_u \cap S_v$ , προσθέτουμε την ακμή  $(Z_{u,g}, Z_{v,g})$  στο γράφημα  $F'$ . Για κάθε κορυφή  $u$  και κάθε μέσο ασύρματης διασύνδεσης  $g \in S_u$ , προσθέτουμε την ακμή  $(Z_{u,0}, Z_{u,g})$  στο  $F'$ . Με παρόμοια επιχειρήματα όπως προηγουμένως, καταλήγουμε στο ότι το γράφημα  $F'$  διατηρεί την απαίτηση συνεκτικότητας μεταξύ κορυφών του ίδιου τερματικού συνόλου  $D'_i$  και στο ότι το συνολικό βάρος των κορυφών του ισούται με το συνολικό κόστος ενεργοποίησης στη βέλτιστη λύση για το στιγμιότυπο  $J$ . Επομένως, το κόστος της βέλτιστης λύσης για το στιγμιότυπο  $J_{NWSF}$  δεν ξεπερνά το κόστος της  $F'$ .  $\square$

Στην εργασία [41], οι Guha και Khuller παρουσιάζουν έναν  $1.61 \ln k$ -προσεγγιστικό αλγόριθμο για το πρόβλημα εύρεσης Δάσους Steiner με βάρη στις κορυφές, όπου  $k$  είναι ο αριθμός των τερματικών στο γράφημα. Χρησιμοποιούμε τον αλγόριθμο αυτόν για να λύσουμε το στιγμιότυπο  $J_{NWSF}$ , και λαμβάνοντας υπόψη και την παραπάνω συζήτηση επιτυγχάνουμε μια λύση

για το στιγμιότυπο  $J$  που είναι στα πλαίσια του  $1.61 \ln |D|$  από τη βέλτιστη. Επομένως, έχουμε:

**Θεώρημα 20.** Υπάρχει ένας  $1.61 \ln |D|$ -προσεγγιστικός αλγόριθμος για τη μη συμμετρική περίπτωση του προβλήματος ομαδικής επικοινωνίας με χρήση πολλαπλασίων μέσω ασύρματης διασύνδεσης GroupMI, όπου  $D$  είναι το σύνολο των τερματικών.

Επιπλέον, αν  $J$  είναι ένα στιγμιότυπο του προβλήματος ConMI, τότε  $p = 1$  και το στιγμιότυπο  $J_{NEWSF}$  αποτελεί ουσιαστικά στιγμιότυπο του προβλήματος εύρεσης δένδρου Steiner με βάρη στις κορυφές το οποίο μπορεί να προσεγγιστεί στα πλαίσια ενός παράγοντα  $1.35 \ln k$ .

**Θεώρημα 21.** Υπάρχει ένας  $1.35 \ln n$ -προσεγγιστικός αλγόριθμος για τη μη συμμετρική περίπτωση του προβλήματος διασύνδεσης με χρήση πολλαπλασίων μέσω ασύρματης διασύνδεσης ConMI σε ασύρματα δίκτυα με  $n$  κόμβους.



## Κεφάλαιο 5

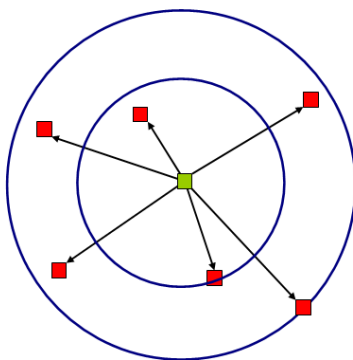
# Ενεργειακά αποδοτική διάδοση μηνυμάτων σε ασύρματα αδόμητα δίκτυα

### 5.1 Εισαγωγή

Τις τελευταίες δεκαετίες τα ασύρματα δίκτυα έχουν αναπτυχθεί ραγδαία. Ειδικότερα, τα ασύρματα δίκτυα τύπου *ad hoc* αναπτύχθηκαν λόγω των ενδεχόμενων εφαρμογών τους σε στρατιωτικές επιχειρήσεις, στην αντιμετώπιση έκτακτων καταστροφών, κτλ. [79]β. Σε αντίθεση με τα παραδοσιακά κυψελικά δίκτυα, στα ασύρματα αδόμητα δίκτυα (ή αλλιώς ασύρματα δίκτυα τύπου *ad hoc*) δεν υπάρχει ενσύρματη δικτυακή υποδομή για την υποστήριξη της ασύρματης επικοινωνίας.

Κάθε κόμβος (ή σταθμός) ενός τέτοιου δικτύου είναι εφοδιασμένος με μια κεραία μέσω της οποίας γίνεται αποστολή και λήψη σημάτων. Η επικοινωνία αποκαθίσταται με την ανάθεση μιας ισχύος μετάδοσης σε κάθε κόμβο. Στο πιο γνωστό μοντέλο εξασθένισης ισχύος [69], η ισχύς του σήματος εξασθενίζει με βάση τη σχέση  $1/r^\alpha$ , όπου  $r$  είναι η απόσταση από τον εκπομπό και  $\alpha$  είναι μια σταθερά που εξαρτάται από το περιβάλλον που βρίσκεται το ασύρματο δίκτυο (συνηθισμένες τιμές για το  $\alpha$  κυμαίνονται μεταξύ 1 και 6). Επομένως, ο εκπομπός  $s$  μπορεί να αποστείλει ένα σήμα στον αποδέκτη  $t$  αν  $P_s \geq \gamma \cdot d(s, t)^\alpha$ , όπου  $P_s$  είναι η ισχύς του μεταδιδόμενου σήματος,  $d(s, t)$  είναι η Ευκλείδεια απόσταση μεταξύ εκπομπού και αποδέκτη, και  $\gamma$  είναι η ελάχιστη ισχύς που πρέπει να έχει ο αποδέκτης προκειμένου να μπορεί να

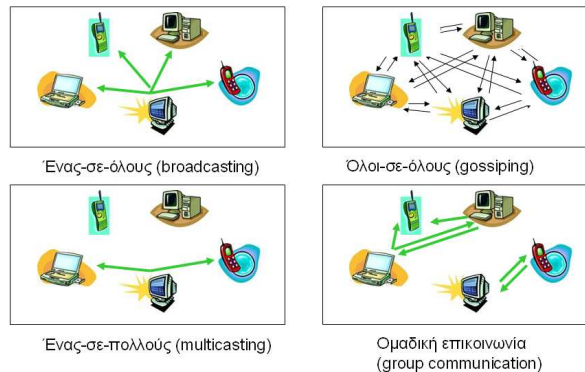
ανιχνεύσει το σήμα και συνήθως κανονικοποιείται στην τιμή 1. Τότε, λέμε ότι ο κόμβος  $s$  εγκαθιστά μια άμεση σύνδεση προς τον κόμβο  $t$  (βλ. και Σχήμα 5.1). Επομένως, επικοινωνία ενός κόμβου  $s$  με έναν άλλο κόμβο  $t$  μπορεί να επιτευχθεί είτε άμεσα αν οι δύο κόμβοι είναι αρκετά κοντά και ο  $s$  χρησιμοποιεί επαρκή ισχύ μετάδοσης, είτε με χρήση ενδιάμεσων κόμβων. Εξαιτίας της μη γραμμικής εξασθένισης της ισχύος, η αποστολή ενός σήματος μεταξύ ενδιάμεσων κόμβων μπορεί να οδηγήσει σε εξοικονόμηση ενέργειας (χρησιμοποιούμε τους όρους ενέργεια και ισχύς, εναλλακτικά).



Σχήμα 5.1: Επικοινωνία: η απαιτούμενη ισχύς ανάλογη του  $r^\alpha$ .

Ένα πολύ σημαντικό ζήτημα στα ασύρματα αδόμητα δίκτυα είναι η υποστήριξη σχημάτων επικοινωνίας που είναι βασικά στα παραδοσιακά δίκτυα, όπως διάδοση (broadcasting), μετάδοση (multicasting), επικοινωνία όλων-προς-όλους (gossiping) και άλλα (Σχ. 5.2). Επειδή η εγκαθίδρυση ενός σχήματος επικοινωνίας εξαρτάται σημαντικά από τη χρήση της ενέργειας, το σημαντικό πρόβλημα που καλούμαστε να αντιμετωπίσουμε είναι να εγυνηθούμε ένα επιθυμητό σχήμα επικοινωνίας ελαχιστοποιώντας τη συνολική κατανάλωση ενέργειας.

Ένα ασύρματο αδόμητο δίκτυο αναπαρίσταται από ένα πλήρες κατευθυνόμενο γράφημα  $G = (V, E)$ , όπου  $|V| = n$  και μια μη αρνητική συνάρτηση κόστους ορισμένη στις ακμές  $c : E \rightarrow R^+$ . Διαισθητικά,  $V$  είναι το σύνολο των σταθμών ή κόμβων, οι ακμές στο σύνολο  $E$  αντιστοιχούν σε ενδεχόμενες άμεσες συνδέσεις, και η συνάρτηση  $c$  συμβολίζει την ελάχιστη ενέργεια που απαιτείται για την εγκαθίδρυση άμεσης σύνδεσης μεταξύ οποιουδήποτε πιθανού ζεύγους αποστολέα-παραλήπτη. Συνήθως, η συνάρτηση κόστους στις ακμές είναι συμμετρική (δηλαδή,  $c(u, v) = c(v, u)$ ).



Σχήμα 5.2: Δημοφιλή σχήματα επικοινωνίας σε ασύρματα αδόμητα δίκτυα.

Μια σημαντική ειδική περίπτωση, η οποία αντανακλά συνήθως την πραγματικότητα, και εφεξής καλείται γεωμετρική περίπτωση, προκύπτει όταν οι κορυφές του γραφήματος  $G$  είναι σημεία στο Ευκλείδειο επίπεδο και το κόστος μιας ακμής  $(u, v)$  ορίζεται ως η Ευκλείδεια απόσταση μεταξύ των κορυφών  $u$  και  $v$  υψωμένη σε μια σταθερή δύναμη  $\alpha$ , δηλ.  $c(u, v) = d(u, v)^\alpha$ . Μη συμμετρικές συναρτήσεις κόστους στις ακμές μπορούν να χρησιμοποιηθούν για την αναπαράσταση της μη κανονικότητας του μέσου μετάδοσης ή μπαταριών με διαφορετικά ενεργειακά επίπεδα [61].

Θεωρούμε ένα δίκτυο που συμβολίζεται με ένα γράφημα  $H = (U, A)$ , όπου  $U \subseteq V$  και  $A \subseteq E$ . Για να κατασκευαστεί το δίκτυο  $H$  πάνω στο αδόμητο ασύρματο δίκτυο  $G$ , πρέπει να θέσουμε το ενεργειακό επίπεδο του κόμβου  $u$  στην τιμή  $\max_{v \in U: (u,v) \in A} c(u, v)$ . Εναλλακτικά, το ενεργειακό επίπεδο στο οποίο ο κόμβος  $u$  λειτουργεί πρέπει να είναι τέτοιο που να του δίνει τη δυνατότητα να εγκαθιδρύσει απευθείας συνδέσεις προς όλες τις ακμές του συνόλου  $A$  που εξέρχονται του κόμβου  $u$ . Η συνολική ενέργεια που απαιτείται για την εγκαθίδρυση του γραφήματος  $H$  πάνω στο  $G$  είναι το άθροισμα των ενεργειακών επιπέδων όλων των κόμβων.

Το πρόβλημα βελτιστοποίησης που μελετάμε είναι το εξής. Δεδομένου ενός αδόμητου ασύρματου δικτύου που αναπαρίσταται από ένα γράφημα  $G = (V, E)$ , με μια μη αρνητική συνάρτηση κόστους στις ακμές  $c : E \rightarrow R^+$ , μια ειδική κορυφή  $r \in V$  που καλείται *ρίζα*, και ένα σύνολο *τερματικών*  $D \subseteq V - \{r\}$ , ζητείται να βρεθεί ένα δένδρο μετάδοσης, δηλ., ένα δένδρο με ρίζα την κορυφή  $r$  που να περιέχει όλες τις κορυφές του συνόλου  $D$ , που να μπορεί να εγκαθιδρυθεί σαν ένα δίκτυο στο  $G$  με την ελάχιστη συνολική

ενέργεια. Το πρόβλημα αυτό είναι γνωστό ως εύρεση ΔΕΝΔΡΟΥ ΜΕΤΑΔΟΣΗΣ ΕΛΑΧΙΣΤΗΣ ΕΝΕΡΓΕΙΑΣ (MINIMUM ENERGY MULTICAST TREE (MEMT)). Η ειδική περίπτωση του προβλήματος MEMT στην οποία  $D = V - \{r\}$  είναι γνωστή ως εύρεση ΔΕΝΔΡΟΥ ΔΙΑΔΟΣΗΣ ΕΛΑΧΙΣΤΗΣ ΕΝΕΡΓΕΙΑΣ (MINIMUM ENERGY BROADCAST TREE (MEBT)).

Στην περίπτωση συμμετρικών συναρτήσεων κόστους στις ακμές, οι Caragiannis et al. [17] παρουσιάζουν λογαριθμικούς (ως προς τον αριθμό  $n$  των σταθμών) προσεγγιστικούς αλγόριθμους για τα προβλήματα MEMT και MEBT. Τα αποτελέσματα αυτά είναι ασυμπτωτικά βέλτιστα αφού MEBT σε συμμετρικά γραφήματα έχει αποδειχθεί ότι τα προβλήματα είναι μη προσεγγίσιμα στα πλαίσια ενός υπολογιστικού παράγοντα [25]. Στην εργασία [17] αποδεικνύεται επίσης ότι, στην περίπτωση μη συμμετρικών συναρτήσεων κόστους στις ακμές, το πρόβλημα MEMT είναι δύσκολο να προσεγγιστεί στα πλαίσια ενός παράγοντα  $O(\log^{2-\epsilon} n)$ , ενώ ο Liang [61] παρουσιάζει έναν  $O(|D|^\epsilon)$  προσεγγιστικό αλγόριθμο. Για το πρόβλημα MEBT στην περίπτωση μη συμμετρικών συναρτήσεων κόστους στις ακμές, έχουν παρουσιασθεί λογαριθμικοί προσεγγιστικοί αλγόριθμοι στις εργασίες [13, 17]. Ο αλγόριθμος στην εργασία [17] για τη συμμετρική περίπτωση του προβλήματος MEMT και ο αλγόριθμος στην εργασία [13] για τη μη συμμετρική περίπτωση του προβλήματος MEBT δανείζονται ιδέες από αλγόριθμους για ένα γνωστό συνδυαστικό πρόβλημα που είναι γνωστό ως πρόβλημα εύρεσης ΔΕΝΔΡΟΥ STEINER ΜΕ ΒΑΡΗ ΣΤΙΣ ΚΟΡΥΦΕΣ (NODE WEIGHTED STEINER TREE (NWST) [41, 56].

Αν και οι παραπάνω περιπτώσεις είναι πολύ ενδιαφέρουσες από θεωρητικής πλευράς, οι πιο σημαντικές ερωτήσεις στην πράξη αφορούν στη γεωμετρική περίπτωση. Η περίπτωση αυτή πρωτομελετήθηκε στην εργασία [54] σε ελαφρώς διαφορετικό πλαίσιο. Οι γεωμετρικές περιπτώσεις του προβλήματος MEBT έχουν λάβει εξαιρετική προσοχή στη βιβλιογραφία. Όταν οι κόμβοι είναι σημεία σε μια γραμμή, το πρόβλημα MEBT μπορεί να λυθεί βέλτιστα σε πολυωνυμικό χρόνο [16, 26]. Η περίπτωση που οι κόμβοι είναι σημεία στο Ευκλείδειο επίπεδο έχει μελετηθεί εκτενώς. Στην περίπτωση αυτή, το πρόβλημα MEBT αποδείχθηκε ότι είναι υπολογιστικά (NP-hard) στην εργασία [25]. Οι πρώτοι αλγόριθμοι προτάθηκαν στην εργασία του Wieselthier και άλλων [79]. Οι αλγόριθμοι αυτοί βασίζονταν στην κατασκευή ελάχιστων γεννητικών δένδρων (MST) και δένδρων ελαχίστων μονοπατιών (SPT) στο γράφημα που αναπαριστά το αδόμητο ασύρματο δίκτυο. Η προσέγγιση που υιοθετήθηκε στην εργασία [79] για τον υπολογισμό λύσεων του προβλήματος MEMT ήταν το 'κλάδεμα' (pruning) των δένδρων που επιστρέφονταν σε λύσεις του προβλήματος MEBT. Πειραματικά αποτελέσματα έδειξαν ότι ο αλ-

γόριθμος *Διάδοσης με Επαυξητική Ενέργεια (Broadcast Incremental Power (BIP))* επιτυγχάνει καλύτερη απόδοση από τους αλγόριθμους εύρεσης ελάχιστων γεννητικών δένδρων (MST) και δένδρων ελαχίστων μονοπατιών (SPT). Σε μετέπειτα δουλειά, οι Wan και άλλοι [78] μελετούν τους αλγόριθμους της εργασίας [79] ως προς την αποδοτικότητα με την οποία προσεγγίζουν τη βέλτιστη λύση. Το βασικό τους αποτέλεσμα είναι ένα άνω φράγμα με τιμή 12 για το λόγο προσέγγισης του αλγόριθμου MST. Το αποτέλεσμα αυτό συνεπάγεται ένα αλγόριθμο σταθερής προσέγγισης και για το πρόβλημα MEMT. Ελαφρώς χειρότερα φράγμα προσέγγισης για το πρόβλημα MST παρουσιάστηκαν στην εργασία [25]. Στην εργασία [78], αποδεικνύεται επίσης ότι ο λόγος προσέγγισης του αλγορίθμου BIP δεν είναι χειρότερος από αυτόν του αλγορίθμου MST, και ότι οι άλλοι διαισθητικοί αλγόριθμοι έχουν κακό λόγο προσέγγισης.

Πάντως, έχει αποδειχθεί πειραματικά ότι αρκετοί διαισθητικοί αλγόριθμοι επιτυγχάνουν καλή απόδοση σε τυχαία στιγμιότυπα του προβλήματος. Στις εργασίες [64, 77], αλγόριθμοι βασισμένοι στην εύρεση ελάχιστων μονοπατιών βελτιώνονται με την ιδέα *ενδεχόμενης εξοικονόμησης ενέργειας (potential power saving idea)* και αποδεικνύεται πειραματικά ότι επιτυγχάνουν καλύτερη απόδοση από τους περισσότερους γνωστούς αλγόριθμους. Στην εργασία [12], ο Cagalj και άλλοι εισήγαγαν μια ευριστική μέθοδο, που είναι γνωστή ως *Ενσωματωμένο Πλεονέκτημα Ασύρματης Μετάδοσης (Embedded Wireless Multicast Advantage-EWMA)* για τον υπολογισμό αποδοτικών λύσεων για για στιγμιότυπα του προβλήματος MEMT. Ο αλγόριθμος αυτός λαμβάνει σαν είσοδο ένα γεννητικό δένδρο και το μετατρέπει σε ένα δένδρο διάδοσης αποδοτικό ως προς την ενέργεια εκτελώντας τοπικές βελτιώσεις. Όπως αναφέρουμε στην Ενότητα 5.2, ο αλγόριθμος EWMA μπορεί εύκολα να τροποποιηθεί ώστε να δουλεύει και για το πρόβλημα MEMT. Μια άλλη ευριστική τεχνική γνωστή με τον όρο Σάρωση (Sweep) παρουσιάστηκε στην εργασία [79]. Ο αλγόριθμος αυτός επίσης λαμβάνει σαν είσοδο ένα δένδρο και το μετατρέπει σε ένα δένδρο αποδοτικό ως προς την ενέργεια εκτελώντας τοπικές βελτιώσεις. Αντίθετα με αυτούς τους δύο αλγόριθμους, οι περισσότεροι από τους προηγούμενους αλγόριθμους βασίζονται στην ιδέα σταδιακής κατασκευής ενός δένδρου. Αυτό σημαίνει ότι, ξεκινώντας από μια κενή λύση, ένα δένδρο επαυξάνεται συμπεριλαμβάνοντας επαναληπτικά νέες δομές (δηλ., νέους κόμβους και ακμές) έως ότου εγκαθιδρυθεί συνεκτικότητα μεταξύ ρίζας και τερματικών. Ένα άλλο θέμα προφανούς σημασίας είναι ο σχεδιασμός αλγορίθμων για το πρόβλημα MEMT που να μπορούν να υλοποιηθούν σε κατανεμημένο περιβάλλον (δείτε για παράδειγμα τις εργασίες [12, 19, 77]).

Στη συνέχεια, χωρίζουμε τους αλγόριθμους της βιβλιογραφίας σε δύο

κατηγορίες: *αλγόριθμους τοπικής αναζήτησης (local search algorithms)* και *επαυξητικούς αλγόριθμους (augmentation algorithms)*. Περιγράφουμε τα γενικά χαρακτηριστικά και των δύο κατηγοριών και παρέχουμε υλοποιήσεις πολλών αλγορίθμων, υπαρχόντων και νέων. Οι υλοποιήσεις μας περιλαμβάνουν αλγόριθμους που προσεγγίζουν το πρόβλημα MEMT στη γενικότερη συμμετρική περίπτωση καθώς και αλγόριθμους που είναι πιο διαισθητικοί για τη γεωμετρική περίπτωση. Στόχος μας είναι η πειραματική σύγκριση των αλγορίθμων αυτών όσον αφορά στην αποδοτικότητα ως προς την ενέργεια των λύσεων που επιτυγχάνουν για τυχαία στιγμιότυπα του προβλήματος MEMT στο Ευκλείδειο επίπεδο. Παρουσιάζουμε επίσης μια αξιολόγηση των χρόνων εκτέλεσης των αλγορίθμων. Η Ενότητα 5.2 αφιερώνεται στους αλγόριθμους τοπικής αναζήτησης, ενώ οι επαυξητικοί αλγόριθμοι παρουσιάζονται στην Ενότητα 5.3. Παρουσιάζουμε και σχολιάζουμε τα πειραματικά αποτελέσματα στην Ενότητα 5.4. Τα αποτελέσματα αυτού του κεφαλαίου έχουν παρουσιαστεί στην εργασία [5].

## 5.2 Αλγόριθμοι τοπικής αναζήτησης

Οι αλγόριθμοι τοπικής αναζήτησης εκτελούν ‘περίπατο’ σε δένδρα μετάδοσης. Ο περίπατος ξεκινά από ένα δένδρο μετάδοσης που δίνεται σαν είσοδος. Σε κάθε βήμα, ένας αλγόριθμος τοπικής αναζήτησης περνάει σε ένα νέο δένδρο μετάδοσης που προκύπτει με την απομάκρυνση κάποιων ακμών από το προηγούμενο δένδρο και την πρόσθεση νέων ακμών, έτσι ώστε να πληρούνται οι αναγκαίες ιδιότητες συνεκτικότητας. Ο κανόνας που χρησιμοποιείται σε κάθε πέρασμα για την επιλογή του επόμενου δένδρου μετάδοσης σχετίζεται με την ενέργεια. Επειδή οι αλγόριθμοι τοπικής αναζήτησης απαιτούν ένα δένδρο μετάδοσης στο οποίο ξεκινούν τον περίπατο, χρησιμοποιούνται συνήθως μετά από έναν επαυξητικό αλγόριθμο. Χαρακτηριστικοί αντιπρόσωποι αυτής της κατηγορίας είναι οι αλγόριθμοι Prune, EWMA και Sweep.

Ο αλγόριθμος ‘κλαδέματος’ Prune έχει χρησιμοποιηθεί εκτενώς (δείτε για παράδειγμα τις εργασίες [64, 79]) για την κατασκευή ενός δένδρου μετάδοσης από ένα δένδρο διάδοσης. Σε κάθε βήμα ο αλγόριθμος εκτελεί την ακόλουθη λειτουργία. Απομακρύνει από το δένδρο κάθε φύλλο που δεν είναι τερματικό μαζί με την εισερχόμενη σε αυτό ακμή. Ο αλγόριθμος τερματίζει όταν όλα τα φύλλα είναι τερματικά. Ο αλγόριθμος Prune μπορεί εύκολα να υλοποιηθεί ώστε να εκτελείται σε γραμμικό χρόνο.

Ο αλγόριθμος EWMA προτάθηκε στην εργασία [12] για τη λύση του προ-

βλήματος ΜΕΒΤ, όπου έγινε η υπόθεση ότι το δένδρο με το οποίο ξεκινά ο αλγόριθμος είναι ένα ελάχιστο γεννητικό δένδρο. Πάντως, όπως είναι προφανές στην περιγραφή του αλγορίθμου που ακολουθεί, μπορεί να χρησιμοποιηθεί για το πρόβλημα ΜΕΜΤ και μπορεί να ξεκινάει με οποιοδήποτε δένδρο μετάδοσης. Ξεκινώντας από ένα δένδρο μετάδοσης, ο αλγόριθμος ΕWMA εκτελεί περιπάτους σε δένδρα μετάδοσης εκτελώντας τους ακόλουθους δύο τύπους αλλαγών σε κάθε βήμα: (1) προστίθενται εξερχόμενες ακμές σε κάποιον κόμβο  $v$ . Ο κόμβος αυτός λέμε ότι *επεκτείνεται* και (2) απομακρύνονται όλες οι εξερχόμενες ακμές από κάποιους απογόνους του  $v$ . Στην περίπτωση αυτή λέμε ότι οι συγκεκριμένοι απόγονοι του  $v$  *αποκλείονται*. Κατά την εκτέλεσή του, ο αλγόριθμος ΕWMA χρησιμοποιεί τρία σύνολα  $C$ ,  $F$  και  $E$ . Διασηθητικά, το  $C$  είναι το σύνολο των κόμβων που έχει λάβει υπόψη του ο αλγόριθμος, το  $F$  είναι το σύνολο των κόμβων που επεκτάθηκαν τουλάχιστον μία φορά σε κάποιο προηγούμενο βήμα και ποτέ δεν αποκλείστηκαν, και το  $E$  είναι το σύνολο των κόμβων που αποκλείστηκαν σε κάποιο προηγούμενο βήμα. Αρχικά, ο αλγόριθμος θέτει  $C = \{r\}$  και  $F = E = \emptyset$ . Σε κάθε βήμα, ο αλγόριθμος ΕWMA λαμβάνει σαν είσοδο το δένδρο μετάδοσης που προέκυψε στο προηγούμενο βήμα μαζί με τα σύνολα  $C$ ,  $F$  και  $E$ . Ορίζουμε σαν *κέρδος* ενός κόμβου  $v$  τη μείωση της ενέργειας που επιτυγχάνεται στο δένδρο μετάδοσης αν αποκλειστούν κάποιοι κόμβοι του δένδρου, σαν αντάλλαγμα για την αύξηση της ενέργειας στον κόμβο  $v$  ώστε να μπορεί να δημιουργήσει ακμές προς όλους τους αποκλειόμενους κόμβους και τα παιδιά τους. Αν κανένας κόμβος στο σύνολο  $C - F - E$  δεν έχει θετικό κέρδος, ο κόμβος με τη μικρότερη ενέργεια συμπεριλαμβάνεται στο σύνολο  $F$  και τα παιδιά του συμπεριλαμβάνονται στο σύνολο  $C$ . Διαφορετικά, ο κόμβος με μέγιστο κέρδος συμπεριλαμβάνεται στο σύνολο  $F$ , οι αποκλειόμενοι κόμβοι συμπεριλαμβάνονται στο σύνολο  $E$ , ενώ και οι αποκλειόμενοι κόμβοι και τα παιδιά τους συμπεριλαμβάνονται στο σύνολο  $C$ . Απομακρύνονται οι ακμές των αποκλειόμενων κόμβων προς τα παιδιά τους (αλλαγές τύπου (2)» και δημιουργούνται εξερχόμενες από τον κόμβο  $v$  προς όλους τους αποκλειόμενους κόμβους και τα παιδιά τους (αλλαγές τύπου (1)». Το νέο δένδρο μετάδοσης μαζί με τα ενημερωμένα σύνολα  $C$ ,  $F$  και  $E$  δίνονται σαν είσοδος στο επόμενο βήμα. Ο αλγόριθμος τερματίζει όταν η ρίζα και όλα τα τερματικά περιλαμβάνονται στο σύνολο  $C$ . Η υλοποίηση του ΕWMA που πραγματοποιήσαμε έχει χρόνο εκτέλεσης  $O(n^3)$  στη χειρότερη περίπτωση.

Ο αλγόριθμος Σάρωση (Sweep) προτάθηκε στην εργασία [79] σαν μια απλή ευριστική μέθοδος για βελτίωση των λύσεων για το πρόβλημα ΜΕΒΤ που επιτυγχάνονταν από αλγόριθμους εύρεσης γεννητικών δένδρων και ελάχιστων

μονοπατιών. Προφανώς, ο αλγόριθμος μπορεί να χρησιμοποιηθεί και για δένδρο μετάδοσης. Ο αλγόριθμος Sweep λειτουργεί ως εξής. Αρχικά αναθέτει διαφορετικές ταυτότητες (IDs) (διαδοχικούς ακέραιους  $0, 1, \dots$ ) σε όλους τους κόμβους. Ξεκινώντας από το δένδρο μετάδοσης, προχωράει σε βήματα. Στο  $i$ -ό βήμα, εξετάζει τον κόμβο  $v_i$  με ταυτότητα (ID) ίση με  $i$ . Αν, για κάποιους κόμβους  $v_{i_1}, v_{i_2}, \dots$  που δεν είναι πρόγονοι του  $v_i$ , η ενέργεια του  $v_i$  στο δένδρο μετάδοσης δεν είναι μικρότερη από το κόστος όλων των ακμών από τον  $v_i$  στους  $v_{i_1}, v_{i_2}, \dots$ , ο αλγόριθμος Sweep απομακρύνει τις εισερχόμενες ακμές από τους κόμβους  $v_{i_1}, v_{i_2}, \dots$  και προσθέτει ακμές από τον κόμβο  $v_i$  στους  $v_{i_1}, v_{i_2}, \dots$  στο δένδρο μετάδοσης. Ο αλγόριθμος τερματίζει όταν εξεταστούν όλοι οι κόμβοι. Αν και φαίνονται προφανείς πολλές παραλλαγές του αλγόριθμου Sweep (όπως αναφέρεται και στην εργασία [79]), αναπάντεχα, καμία από τις αρκετές παραλλαγές που υλοποιήσαμε δεν είναι καλύτερη από τον αλγόριθμο Sweep ως προς την αποδοτικότητα ως προς την ενέργεια της επιστρεφόμενης λύσης. Μια απλή υλοποίηση του αλγορίθμου Sweep εκτελείται σε χρόνο  $O(n^2)$ .

### 5.3 Επαυξητικοί αλγόριθμοι

Οι επαυξητικοί αλγόριθμοι κατασκευάζουν δένδρο μετάδοσης ξεκινώντας από μία κενή λύση η οποία σταδιακά επαυξάνεται μέχρι να δημιουργηθεί ένα δίκτυο με κατευθυνόμενα μονοπάτια από τη ρίζα στα τερματικά. Προφανώς, από το στιγμή που ένα τέτοιο δίκτυο είναι διαθέσιμο, μπορεί εύκολα να μετατραπεί σε ένα δένδρο μετάδοσης. Η λύση επαυξάνεται σε φάσεις. Σε κάθε φάση, ένας επαυξητικός αλγόριθμος προσθέτει μια δομή στη λύση. Αυτή μπορεί να είναι μια ακμή σε έναν κόμβο (π.χ., στον αλγόριθμο BIP και σε γνωστές υλοποιήσεις του αλγορίθμου MST [29]), ένα σύνολο ακμών που εξέρχονται συγκεκριμένου κόμβου (π.χ., σε μια παραλλαγή του αλγορίθμου BIP που λέγεται BAIP [79]), ένα μονοπάτι (π.χ., σε αλγόριθμους που παρουσιάζονται στις εργασίες [64, 77]), μία αράχνη (spider) (ένα ειδικό κατευθυνόμενο γράφημα που χρησιμοποιείται στην εργασία [13] και έμμεσα στην εργασία [17]), κτλ. Η δομή επιλέγεται μεταξύ όλων των υποψήφιων δομών έτσι ώστε κάποιος τοπικός στόχος να ελαχιστοποιείται. Ο τοπικός στόχος σχετίζεται συνήθως με την ενέργεια που απαιτείται για τη δημιουργία των ακμών της δομής. Στη συνέχεια παραθέτουμε μια αναλυτική περιγραφή αρκετών επαυξητικών αλγορίθμων, που περιλαμβάνουν υπάρχοντες αλγόριθμους από τη βιβλιογραφία αλλά και νέους και αρκετές παραλλαγές τους.



**Βασικοί αλγόριθμοι.** Ξεκινώντας από ένα δένδρο μετάδοσης που περιέχει μόνο τη ρίζα, οι αλγόριθμοι *Shortest Path First* (SPF), *Multicast Incremental Power* (MIP), *Densest Shortest Path First* (DSPF) και *Densest Two Shortest Path First* (D2SPF) επαυξάνουν το δένδρο μετάδοσης σε φάσεις, μέχρι όλα τα τερματικά να συμπεριληφθούν στο δένδρο. Σε κάθε φάση, ο αλγόριθμος SPF προσθέτει στο δένδρο μετάδοσης το ελάχιστο κατευθυνόμενο μονοπάτι (δηλαδή, το μονοπάτι του οποίου η δημιουργία απαιτεί το ελάχιστο ποσό ενέργειας) που συνδέει κάποιον κόμβο του δένδρου σε κάποιο τερματικό του δένδρου. Ο αλγόριθμος MIP προσθέτει το μονοπάτι που απαιτεί την ελάχιστη επιπλέον ενέργεια που συνδέει το δένδρο με κάποιο τερματικό εκτός του δένδρου. Αυτό σημαίνει ότι μετά από κάθε φάση το κόστος των ακμών που εξέρχονται κάθε κόμβου  $v$  στο μονοπάτι που επιλέγεται κατά τη φάση ελαττώνεται κατά το κόστος της ακμής που εξέρχεται του  $v$  στο μονοπάτι. Ο αλγόριθμος DSPF επιλέγει από τα μονοπάτια του δένδρου που απαιτούν ελάχιστη επιπλέον ενέργεια για σύνδεση με τα τερματικά εκτός του δένδρου, εκείνο που ελαχιστοποιεί το λόγο της επιπλέον ενέργειας προς τον αριθμό των νέων τερματικών που συμπεριλαμβάνονται στο δένδρο. Ο αλγόριθμος D2SPF είναι ίδιος με τον DSPF. Η διαφορά τους είναι ότι για κάθε κόμβο  $v$  και κάθε τερματικό  $t$  που και δεν ανήκουν στο δένδρο μετάδοσης, ο αλγόριθμος D2SPF λαμβάνει υπόψη του μονοπάτια από το δένδρο στο τερματικό  $t$  που περνούν από τον κόμβο  $v$ .

Στις υλοποιήσεις μα, χρησιμοποιούμε εκτενώς αλγόριθμους υπολογισμού ελάχιστων μονοπατιών. Ο υπολογισμός ελάχιστων μονοπατιών από έναν κόμβο προς όλους τους άλλους σε ένα πλήρες γράφημα μπορεί να πραγματοποιηθεί σε χρόνο  $O(n^2)$  χρησιμοποιώντας μια απλή υλοποίηση του αλγορίθμου του Dijkstra με χρήση δυαδικών σωρών για ουρές προτεραιότητας [29]. Πιο πολύπλοκες σωροί που έχει αποδειχθεί ότι δίνουν ταχύτερες υλοποιήσεις του αλγορίθμου του Dijkstra (π.χ., σωροί Fibonacci) δε μειώνουν το χρόνο εκτέλεσης σημαντικά στη δική μας περίπτωση αφού το γράφημα που αναπαριστά το αδόμητο ασύρματο δίκτυο είναι πλήρες. Οι αλγόριθμοι SPF, MIP και DSPF εκτελούνται σε χρόνο  $O(mn^2)$ , όπου  $m$  είναι ο αριθμός των τερματικών. Σε κάθε μία από τις το πολύ  $m$  φάσεις, οι αλγόριθμοι εκτελούν έναν υπολογισμό ελάχιστου μονοπατιού. Η επεξεργασία του γραφήματος (δηλαδή, η μείωση στα κόστη των ακμών) που απαιτείται από τους αλγόριθμους MIP και DSPF δεν επηρεάζει τους ασυμπτωτικούς χρόνους εκτέλεσής τους σε σύγκριση με τον αλγόριθμο SPF. Κάθε φάση του αλγορίθμου D2SPF απαιτεί τον υπολογισμό όλων των ζευγών ελαχίστων μονοπατιών που χρειάζεται χρόνο  $O(n^3)$  δίνοντας συνολικό χρόνο εκτέλεσης  $O(mn^3)$ .

**Η ιδέα της ενδεχόμενης εξοικονόμησης ενέργειας.** Οι βασικοί αλγόριθμοι δεν εξετάζουν αν η εγκαθίδρυση ενός νέου μονοπατιού μπορεί επίσης να περιλαμβάνει κόμβους που συμπεριλήφθηκαν στο δένδρο μετάδοσης σε προηγούμενες φάσεις, και τώρα θα μπορούσαν να συνδεθούν στο δένδρο μετάδοσης σαν παιδιά κάποιου κόμβου στο μονοπάτι. Με τον τρόπο αυτό, η ενέργεια του προηγούμενου γονέα τους θα μπορούσε να μειωθεί. Η συνολική μείωση εκφέρεται με τον όρο *ενδεχόμενη εξοικονόμηση ενέργειας* (*potential power saving*) ([64, 77]). Οι αλγόριθμοι που περιγράφουμε είναι παραλλαγές των βασικών αλγορίθμων με τη διαφορά ότι όταν υπολογίζουν την επιπλέον ενέργεια ενός υποψήφιου μονοπατιού, αφαιρούν την ενδεχόμενη εξοικονόμηση ενέργειας. Όταν δημιουργείται ένα νέο μονοπάτι, το δένδρο μετάδοσης τροποποιείται ανάλογα, δηλαδή, κόμβοι που συμπεριλαμβάνονταν στο δένδρο μετάδοσης μπορεί πλέον να συνδέονται σαν παιδιά κόμβων του νέου μονοπατιού.

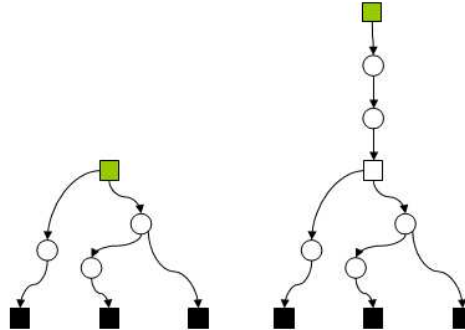
Οι αλγόριθμοι SP3SF, DSP3SF και DSP3SF είναι παραλλαγές των αλγορίθμων MIP, DSPF και D2SPF κάνοντας χρήση της ιδέας ενδεχόμενης εξοικονόμησης ενέργειας, αντίστοιχα. Ο αλγόριθμος SP3SF προτάθηκε αρχικά στην εργασία [64] και, με βάση όσα γνωρίζουμε, υπολογίζει τις πιο αποδοτικές ενεργειακά λύσεις για το πλαίσιο που μελετάμε. Ο αλγόριθμος 2SP3SF είναι παραλλαγή του αλγορίθμου D2SP3SF όπου ο τοπικός στόχος σε κάθε φάση είναι η ελαχιστοποίηση της επιπλέον ενέργειας μείον την ενδεχόμενη εξοικονόμηση ενέργειας. Οι αναγκαίοι υπολογισμοί για τον υπολογισμό της ενδεχόμενης εξοικονόμησης ενέργειας αυξάνουν το χρόνο εκτέλεσης των αλγορίθμων SP3SF και DSP3SF σε  $O(m^2n^2)$ , ενώ οι αλγόριθμοι 2SP3SF και D2SP3SF εκτελούνται σε χρόνο  $O(mn^3)$ . Όμως, στην πράξη ο χρόνος εκτέλεσης των αλγορίθμων SP3SF και DSP3SF είναι ελάχιστα χειρότερος από αυτόν των MIP και DSPF, αντίστοιχα.

**Αλγόριθμοι βασισμένοι το πρόβλημα εύρεσης δένδρου Steiner με βάρη στις κορυφές (NWST).** Παρουσιάζουμε στη συνέχεια έξι αλγόριθμους που δανείζονται ιδέες από προσεγγιστικούς αλγόριθμους για το πρόβλημα εύρεσης δένδρου Steiner με βάρη στις κορυφές (NWST) που παρουσιάστηκαν στις εργασίες [41, 56]. Στην πραγματικότητα, ο αλγόριθμος της εργασίας [17] για τη συμμετρική περίπτωση του προβλήματος MEMT ανάγει στιγμιότυπα του προβλήματος MEMT σε στιγμιότυπα του προβλήματος NWST. Αν και, το μέγεθος του στιγμιότυπου του προβλήματος NWST που προκύπτει είναι πολυωνυμικό με βάση τη θεωρία, ακόμα και για μικρά στιγμιότυπα του MEMT το αντίστοιχο στιγμιότυπο του προβλήματος NWST δε μπορεί να λυθεί με τους αλγόριθμους των εργασιών [41, 56] στην πράξη. Οι αλγόριθμοι

που παρουσιάζονται στη συνέχεια εφαρμόζουν ιδέες από τις εργασίες [41, 56] κατευθείαν στο στιγμιότυπο του προβλήματος MEMT. Ένας παρόμοιος αλγόριθμος έχει σχεδιασθεί για το πρόβλημα MEBT στην εργασία [13]. Όλοι αυτοί οι αλγόριθμοι χρησιμοποιούν την ιδέα σταδιακής επαύξησης ενός δικτύου μέσω επαναλαμβανόμενης πρόσθεσης αραχών (spiders) ή πηρουνιών (forks) μικρότερης πυκνότητας.

Μια *αράχνη* (spider) είναι ένα κατευθυνόμενο γράφημα που αποτελείται από έναν κόμβο που καλείται *κεφαλή* (head) και από ένα σύνολο κατευθυνόμενων μονοπατιών που καλούνται *βραχίονες* (legs), καθένα από τα οποία κατευθύνεται από την κεφαλή στους κόμβους που καλούνται *πόδια* (feet) της αράχνης (βλ. αριστερό μέρος Σχ. 5.3). Ο ορισμός επιτρέπει στους βραχίονες να έχουν κοινούς κόμβους και ακμές. Το βάρος της αράχνης είναι το μέγιστο κόστος των ακμών που εξέρχονται της κεφαλής συν το άθροισμα του επιμέρους κόστους των βραχιόνων, όπου κόστος ενός βραχίονα είναι το άθροισμα του επιμέρους κόστους των ακμών του εκτός από την ακμή που εξέρχεται της κεφαλής. Ένα *πηρούνι* (fork) είναι μια αράχνη που διαθέτει έναν κόμβο *κέντρο* (center) ο οποίος είναι προσπελάσιμος μέσω ενός ελάχιστου κατευθυνόμενου μονοπατιού από την κεφαλή έτσι ώστε το υπογράφημα του πηρουνιού που προκύπτει αν αφαιρέσουμε τις ακμές αυτού του μονοπατιού και όλους τους κόμβους στο μονοπάτι εκτός από το κέντρο να είναι μια αράχνη με τον κόμβο κέντρο σαν κεφαλή της (βλ. δεξί μέρος Σχ. 5.3). Καλούμε αυτή την αράχνη *υποαράχνη* (subspider) του πηρουνιού. Το βάρος του πηρουνιού είναι το κόστος των ακμών στο κατευθυνόμενο μονοπάτι από την κεφαλή στο κέντρο συν το βάρος της υπο-αράχνης του. Λέμε ότι ένας κόμβος  $u$  μπορεί να συνδεθεί *χωρίς κόστος* σε ένα μονοπάτι, αν το κόστος  $c(v_i, u)$  της ακμής από έναν κόμβο  $v_i$  στο μονοπάτι προς τον κόμβο  $u$  είναι μικρότερο από το κόστος της ακμής που εξέρχεται του κόμβου  $v_i$  στο μονοπάτι.

Ο αλγόριθμος *Densest Spider First* (DSF) κατασκευάζει ένα δίκτυο  $H$  στο οποίο η ρίζα και τα τερματικά περιέχονται στην ίδια χαλαρά συνεκτική συνιστώσα. Χαλαρή συνεκτικότητα ενός κατευθυνόμενου γραφήματος σημαίνει ότι το μη κατευθυνόμενο ισοδύναμό του γράφημα (δηλαδή, το γράφημα που λαμβάνεται αντικαθιστώντας τις κατευθυνόμενες ακμές με μη κατευθυνόμενες) είναι συνεκτικό. Τότε, για τη δημιουργία ενός δένδρου μετάδοσης στο γράφημα  $G$ , ο αλγόριθμος υπολογίζει ένα δένδρο με κατεύθυνση από τη ρίζα στα τερματικά στο υπεργράφημα του  $H$  που περιέχει όλες τις ακμές του  $H$  και τις αντίθετα κατευθυνόμενες ακμές τους. Ο αλγόριθμος αναθέτει δείκτες τους κόμβους για να θυμάται ποιοι κόμβοι ανήκουν στην ίδια χαλαρά συνεκτική συνιστώσα του γραφήματος  $H$ . Κάθε δείκτης είναι ένας μη αρ-



Σχήμα 5.3: Αράχνη (spider) και πηρούνι (fork).

νητικός ακέραιος. Οι κόμβοι με τον ίδιο πεπερασμένο δείκτη είναι χαλαρά συνδεδεμένοι και κόμβοι με διαφορετικούς ή άπειρους δείκτες δε συνδέονται καθόλου. Αρχικά, ο δείκτης της ρίζας είναι 0, τα τερματικά έχουν πεπερασμένους, θετικούς και διακριτούς ακέραιους δείκτες, ενώ όλοι οι άλλοι κόμβοι έχουν άπειρο δείκτη. Ο αλγόριθμος προχωράει σε βήματα μέχρι η ρίζα και όλα τα τερματικά να έχουν δείκτη 0. Σε κάθε βήμα, ο αλγόριθμος βρίσκει έναν κόμβο  $v$  και μια αράχνη που έχει τον  $v$  σαν κεφαλή της και κόμβους που ανά δυο έχουν διακριτούς, διαφορετικούς από το  $v$  και μη άπειρους δείκτες για πόδια, έτσι ώστε ο λόγος του βάρους της αράχνης προς τον αριθμό των ποδιών της (που καλείται πυκνότητα της αράχνης) να ελαχιστοποιείται. Έστω  $i_1, i_2, \dots, i_k$  οι δείκτες των κόμβων στην αράχνη με αύξουσα διάταξη. Για κάθε κόμβο στο γράφημα με δείκτη  $i_1$  ή  $i_2, \dots, \text{ ή } i_k$  ο αλγόριθμος θέτει το δείκτη του στην τιμή  $i_1$  και προσθέτει τις ακμές της αράχνης στο  $H$ . Οι αλγόριθμοι DSF-2 και DSF-3 αποτελούν παραλλαγές του αλγορίθμου DSF. Η μόνη διαφορά είναι ότι ο κόμβος που επιλέγεται σαν κεφαλή της αράχνης σε κάθε βήμα πρέπει να έχει πεπερασμένο δείκτη στον αλγόριθμο DSF-2 και δείκτη 0 στον αλγόριθμο DSF-3.

Στις υλοποιήσεις μας, οι αλγόριθμοι αρχικά εκτελούν μια προεπεξεργασία για να υπολογίσουν, για κάθε κόμβο και κάθε πιθανό επίπεδο ενέργειάς του, τα συντομότερα μονοπάτια από τον κόμβο αυτό προς όλους τους άλλους. Αυτό απαιτεί χρόνο  $O(n^4)$  και κυριαρχεί του ασυμπτωτικού χρόνου εκτέλεσης των αλγορίθμων. Αφού υπολογιστεί το μήκος όλων των ελάχιστων μονοπατιών, τότε ο υπολογισμός της αράχνης με τη μικρότερη πυκνότητα που έχει σαν κεφαλή ένα συγκεκριμένο κόμβο με συγκεκριμένο επίπεδο ενέργειας μπορεί

να πραγματοποιηθεί σε χρόνο  $O(m)$ . Επομένως, η αράχνη με τη μικρότερη πυκνότητα σε κάθε μία από τις το πολύ  $m$  φάσεις υπολογίζεται σε χρόνο  $O(mn^2)$ .

Ο αλγόριθμος *Densest Incremental Spider First* (DISF) είναι παραλλαγή του αλγορίθμου DSF-3. Σε κάθε βήμα, αφού επιλεγεί η αράχνη με τη μικρότερη πυκνότητα, για κάθε κόμβο  $u$  στην αράχνη, το κόστος των ακμών που εξέρχονται του  $u$  στο γράφημα  $G$  ελαττώνεται κατά το μέγιστο κόστος των ακμών που εξέρχονται του  $u$  στην αράχνη. Διαισθητικά, το βάρος της αράχνης στους αλγόριθμους DSF, DSF-2 και DSF-3 αποτελεί ένα άνω φράγμα για την ενέργεια που απαιτείται για τη δημιουργία των ακμών της αράχνης. Στον αλγόριθμο DISF, το βάρος της αράχνης που υπολογίζεται σε κάθε βήμα αποτελεί ένα άνω φράγμα για την επιπλέον ενέργεια που απαιτείται για την εγκαθίδρυση των ακμών της αράχνης, δεδομένου ότι ακμές που συμπεριλήφθηκαν στο γράφημα  $H$  σε προηγούμενα βήματα έχουν ήδη εγκαθιδρυθεί. Εξαιτίας της ενημέρωσης της συνάρτησης ανάθεσης βάρους στις ακμές που απαιτείται σε κάθε φάση του αλγορίθμου DISF, απαιτείται σε κάθε φάση παραπλήσια προεπεξεργασία με αυτή του αλγορίθμου DSF. Αυτό απαιτεί χρόνο  $O(mn^4)$  και κυριαρχεί του ασυμπτωτικού χρόνου εκτέλεσης του αλγορίθμου.

Ο αλγόριθμος *Densest Fork First* (DFF) είναι μια παραλλαγή του αλγορίθμου DISF. Σε κάθε βήμα ένα πηρούνι που έχει σαν κεφαλή ένα κόμβο με δείκτη 0 που ελαχιστοποιεί το λόγο του βάρους του πηρουιού προς τον αριθμό των ποδιών προστίθεται στο δίκτυο  $H$ . Τελικά, ο αλγόριθμος DFF-2 είναι ίδιος με τον DFF. Η βασική διαφορά είναι ότι ο αριθμός των ποδιών συν ο αριθμός των μη μηδενικών πεπερασμένων δεικτών των κόμβων που μπορούν να συνδεθούν χωρίς κόστος στο μονοπάτι από την κεφαλή προς το κέντρο χρησιμοποιείται στον παρονομαστή του τοπικού στόχου. Ο χρόνος εκτέλεσης των αλγορίθμων DFF και DFF-2 είναι ασυμπτωτικά ο ίδιος με αυτόν του αλγορίθμου DISF. Σε κάθε φάση ο χρόνος που απαιτείται για προεπεξεργασία ασυμπτωτικά κυριαρχεί του χρόνου που απαιτείται για τον υπολογισμό του πυκνότερου πηρουιού.

**Περιορισμένες και επαναληπτικές εκδοχές.** Σε πολλούς από τους παραπάνω αλγόριθμους, περισσότερα από ένα τερματικά μπορούν να προστεθούν στο δένδρο μετάδοσης σε κάθε φάση. Υπάρχει περίπτωση η πρόσθεση μιας δομής με πολλά τερματικά σε κάποια φάση να χειροτερεύει την τελική λύση. Έχουμε υλοποιήσει περιορισμένες εκδοχές των αλγορίθμων που περιγράφηκαν παραπάνω που λαμβάνουν σαν είσοδο μια *παράμετρο περιορισμού επαύξησης* που συμβολίζει το μέγιστο αριθμό νέων τερματικών που

επιτρέπεται να συμπεριληφθούν σε κάθε φάση και περιορίζει κατάλληλα το χώρο των υποψήφιων δομών. Επαναληπτικές εκδοχές των αλγορίθμων DSPF, DSP3SF, SP3SF, D2SPF, D2SP3SF, 2SP3SF, DISF, DFF και DFF-2 επαναλαμβανόμενα εκτελούν τις περιορισμένες εκδοχές τους για όλες τις πιθανές τιμές της παραμέτρου περιορισμού επαύξησης και επιστρέφουν την καλύτερη λύση. Οι αλγόριθμοι αυτοί καλούνται iDSPF, iDSP3SF, iSP3SF, iD2SPF, iD2SP3SF, i2SP3SF, iDISF, iDFF και iDFF-2, αντίστοιχα. Προφανώς, μια επαναληπτική εκδοχή ενός αλγόριθμου είναι καλύτερη από τη μη περιορισμένη εκδοχή του ως προς την αποδοτικότητά του ως προς την ενέργεια. Όμως, μπορεί να απαιτήσει χρόνο εκτέλεσης ανάλογο προς το γινόμενο του αριθμού των τερματικών επί το χρόνο εκτέλεσης της μη περιορισμένης εκδοχής.

## 5.4 Πειραματικά αποτελέσματα

Στη συνέχεια, παρουσιάζουμε τα αποτελέσματα των πειραμάτων μας με τους αλγόριθμους που παρουσιάσαμε στις προηγούμενες ενότητες. Τα αποτελέσματα των πειραμάτων μας παρουσιάζονται στους Πίνακες 5.1-5.5. Έχουμε επίσης υλοποιήσει μερικούς ακόμα αλγόριθμους διάδοσης (δηλ., παραλλαγές των αλγορίθμων BIP και BAIP) και έχουμε εκτελέσει πειράματα με αυτούς εκτελώντας τον αλγόριθμο Prune στις λύσεις τους για να επιτύχουμε δένδρα μετάδοσης. Τα αποτελέσματα αυτά είναι σαφώς χειρότερα από αυτά που παρουσιάζονται στη συνέχεια. Επιπλέον, έχουμε μελετήσει παραλλαγές των αλγορίθμων που είναι κατάλληλες για υλοποίηση σε κατανεμημένο περιβάλλον. Τα αντίστοιχα πειραματικά αποτελέσματα είναι συνήθως χειρότερα από αυτά των κεντρικοποιημένων αλγορίθμων. Δεν αναφερόμαστε σε θέματα κατανεμημένης υλοποίησης, αλλά εστιάζουμε σε κεντρικοποιημένες υλοποιήσεις αφού παρέχουν κάτω φράγματα για την αποδοτικότητα των λύσεων ως προς την ενέργεια και το συνολικό έργο που απαιτείται για τις κατανεμημένες υλοποιήσεις τους.

Οι αλγόριθμοι εκτελέστηκαν σε διαφορετικού μεγέθους γεωμετρικά στιγμιότυπα του προβλήματος. Τα στιγμιότυπα εισόδου αποτελούνται από κόμβους που αντιστοιχούν σε σημεία με ομοιόμορφα τυχαίες συντεταγμένες στο διάστημα  $[0, 5)$ . Ένας κόμβος επιλέγεται τυχαία να είναι η ρίζα και τα τερματικά επιλέγονται ομοιόμορφα τυχαία χωρίς αντικατάσταση μεταξύ όλων των υπόλοιπων κόμβων. Για κάθε στιγμιότυπο, χρησιμοποιούμε τον όρο *ομάδα* για το σύνολο τερματικών μαζί με τη ρίζα.

Βασικοί αλγόριθμοι και αλγόριθμοι που χρησιμοποιούν την ιδέα ενδεχό-

μενης εξοικονόμησης ενέργειας εκτελέστηκαν σε στιγμιότυπα μεγέθους 100 για ομάδες μεγέθους 10, 20, ..., 100 και τιμές 2 και 4 για την παράμετρο  $\alpha$ . Ο πίνακας 5.1 δείχνει την απόδοση των βασικών αλγορίθμων και των αλγορίθμων που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας ως προς την αποδοτικότητα ως προς την ενέργεια. Για κάθε μέγεθος ομάδας, κατασκευάζονται 100 τυχαία στιγμιότυπα με το συγκεκριμένο μέγεθος ομάδας και εκτελούνται όλοι οι αλγόριθμοι σε αυτά τα τυχαία στιγμιότυπα. Οι τιμές της ενέργειας που φαίνονται στον Πίνακα 5.1 για τις εκτελέσεις κάθε αλγόριθμου σε συγκριμένο μέγεθος ομάδας, είναι η μέση ενέργεια των δένδρων μετάδοσης που υπολογίζονται από τον αλγόριθμο στις εκτελέσεις του σε όλα τα τυχαία στιγμιότυπα με το συγκεκριμένο μέγεθος ομάδας.

Όλοι οι αλγόριθμοι που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας έχουν καλύτερη απόδοση από τους βασικούς αλγόριθμους (αυτό παρατηρήθηκε πειραματικά για τους αλγόριθμους SP3SF και MIP στις εργασίες [64, 77]). Η διαφορά αυτή είναι μεγαλύτερη όταν  $\alpha = 2$ . Έχει ενδιαφέρον ότι, ο αλγόριθμος D2SPF είναι το πολύ 1% χειρότερος από τον αλγόριθμο SP3SF στην περίπτωση αυτή. Ο αλγόριθμος D2SP3SF έχει καλύτερη απόδοση από όλους τους αλγόριθμους που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας όταν  $\alpha = 2$ . Είναι εξαιρετικά καλύτερος από τους αλγόριθμους SP3SF και 2SP3SF και κάπως καλύτερος από τον DSP3SF. Επίσης, ο αλγόριθμος 2SP3SF είναι οριακά καλύτερος από τον SP3SF. Όταν  $\alpha = 4$ , οι αλγόριθμοι SP3SF και 2SP3SF φαίνεται να παράγουν τις πιο αποδοτικές ως προς την ενέργεια λύσεις (με τον αλγόριθμο DSP3SF να είναι οριακά χειρότερος). Στα πειράματά μας, παρατηρήσαμε ότι οι αλγόριθμοι που χρησιμοποιούν δύο ενωμένα ελάχιστα μονοπάτια σε δομές είναι πάντα αρκετά πιο αργοί από αλγόριθμους που χρησιμοποιούν σε δομές απλά ελάχιστα μονοπάτια (δείτε π.χ., τον Πίνακα 5.5). Αυτό είναι ενδιαφέρον (και κάπως αναπάντεχο) αφού ο ασυμπτωτικός χρόνος εκτέλεσης όλων των αλγορίθμων που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας είναι ουσιαστικά ο ίδιος για μεγάλα μεγέθη ομάδων.

Γενικά, οι αλγόριθμοι που βασίζονται σε εύρεση δένδρου Steiner με βάρη στις κορυφές (NWST) είναι αργοί. Αυτό έχει ήδη τεκμηριωθεί στην Ενότητα 5.3 όπου συζητάμε το χρόνο εκτέλεσής τους (δείτε επίσης και τον Πίνακα 5.5). Το γεγονός αυτό δε μάς επέτρεψε να πραγματοποιήσουμε μεγάλα πειράματα. Τα πειράματά μας με στιγμιότυπα του προβλήματος μεγέθους 40 (δείτε τον Πίνακα 5.2) δείχνουν ότι αλγόριθμοι βασισμένοι σε NWST είναι κατώτεροι από τους περισσότερους αλγόριθμους που αναφέραμε προηγουμένως. Οι αλγόριθμοι DSF, DSF-2 και DSF-3 δεν είναι πολύ πιο αργοί από τους

| $\alpha = 2$ | 10   | 20   | 30   | 40   | 50    | 60    | 70    | 80    | 90    | 100   |
|--------------|------|------|------|------|-------|-------|-------|-------|-------|-------|
| SPF          | 5.27 | 7.19 | 8.49 | 9.51 | 10.07 | 10.56 | 11.26 | 11.55 | 12.09 | 12.45 |
| MIP          | 4.98 | 6.78 | 7.96 | 8.85 | 9.47  | 9.89  | 10.52 | 10.84 | 11.26 | 11.61 |
| DSPF         | 4.93 | 6.64 | 7.83 | 8.66 | 9.22  | 9.61  | 10.17 | 10.40 | 10.73 | 11.07 |
| D2SPF        | 5.01 | 6.59 | 7.80 | 8.56 | 9.06  | 9.46  | 10.00 | 10.22 | 10.53 | 10.86 |
| SP3SF        | 4.92 | 6.59 | 7.66 | 8.50 | 9.03  | 9.38  | 9.90  | 10.10 | 10.49 | 10.79 |
| DSP3SF       | 4.87 | 6.51 | 7.59 | 8.43 | 8.88  | 9.25  | 9.76  | 10.01 | 10.26 | 10.63 |
| 2SP3SF       | 4.90 | 6.56 | 7.62 | 8.45 | 8.99  | 9.32  | 9.89  | 10.07 | 10.39 | 10.74 |
| D2SP3SF      | 4.96 | 6.50 | 7.60 | 8.34 | 8.87  | 9.18  | 9.69  | 9.90  | 10.18 | 10.49 |
| IDSPF        | 4.89 | 6.55 | 7.70 | 8.52 | 9.07  | 9.44  | 9.97  | 10.23 | 10.56 | 10.92 |
| ID2SPF       | 4.87 | 6.49 | 7.61 | 8.41 | 8.89  | 9.30  | 9.78  | 10.05 | 10.39 | 10.71 |
| ISP3SF       | 4.92 | 6.59 | 7.66 | 8.50 | 9.03  | 9.38  | 9.90  | 10.10 | 10.48 | 10.79 |
| IDSP3SF      | 4.81 | 6.37 | 7.38 | 8.16 | 8.58  | 8.94  | 9.41  | 9.63  | 9.92  | 10.28 |
| I2SP3SF      | 4.90 | 6.55 | 7.62 | 8.45 | 8.99  | 9.32  | 9.88  | 10.07 | 10.38 | 10.73 |
| ID2SP3SF     | 4.81 | 6.38 | 7.40 | 8.16 | 8.64  | 8.98  | 9.47  | 9.68  | 9.96  | 10.25 |
| $\alpha = 4$ | 10   | 20   | 30   | 40   | 50    | 60    | 70    | 80    | 90    | 100   |
| SPF          | 1.51 | 1.97 | 2.46 | 2.65 | 2.89  | 2.99  | 3.19  | 3.25  | 3.40  | 3.61  |
| MIP          | 1.48 | 1.93 | 2.41 | 2.59 | 2.82  | 2.93  | 3.10  | 3.18  | 3.32  | 3.52  |
| DSPF         | 1.49 | 1.93 | 2.41 | 2.61 | 2.82  | 2.93  | 3.11  | 3.19  | 3.32  | 3.52  |
| D2SPF        | 1.49 | 1.94 | 2.42 | 2.60 | 2.81  | 2.91  | 3.10  | 3.18  | 3.29  | 3.52  |
| SP3SF        | 1.47 | 1.91 | 2.36 | 2.55 | 2.74  | 2.85  | 3.02  | 3.08  | 3.21  | 3.41  |
| DSP3SF       | 1.46 | 1.92 | 2.38 | 2.55 | 2.75  | 2.87  | 3.04  | 3.12  | 3.25  | 3.43  |
| 2SP3SF       | 1.46 | 1.90 | 2.36 | 2.55 | 2.74  | 2.85  | 3.02  | 3.09  | 3.21  | 3.41  |
| D2SP3SF      | 1.48 | 1.92 | 2.38 | 2.55 | 2.75  | 2.86  | 3.03  | 3.12  | 3.22  | 3.41  |
| IDSPF        | 1.47 | 1.91 | 2.38 | 2.56 | 2.78  | 2.89  | 3.06  | 3.13  | 3.27  | 3.46  |
| ID2SPF       | 1.47 | 1.90 | 2.37 | 2.55 | 2.77  | 2.87  | 3.05  | 3.12  | 3.24  | 3.44  |
| ISP3SF       | 1.47 | 1.91 | 2.36 | 2.55 | 2.74  | 2.85  | 3.02  | 3.08  | 3.21  | 3.41  |
| IDSP3SF      | 1.45 | 1.88 | 2.33 | 2.50 | 2.70  | 2.81  | 2.96  | 3.04  | 3.16  | 3.34  |
| I2SP3SF      | 1.46 | 1.90 | 2.36 | 2.55 | 2.74  | 2.85  | 3.02  | 3.09  | 3.21  | 3.41  |
| ID2SP3SF     | 1.46 | 1.89 | 2.34 | 2.51 | 2.71  | 2.82  | 2.99  | 3.06  | 3.17  | 3.35  |

Πίνακας 5.1: Σύγκριση βασικών αλγορίθμων, αλγορίθμων που χρησιμοποιούν την ιδέα της ενδεχόμενης εξοικονόμησης ενέργειας και τις επαναληπτικές τους εκδοχές για τυχαία στιγμιότυπα με 100 κόμβους,  $\alpha = 2$  και  $\alpha = 4$  και μεγέθη ομάδων 10, 20, ..., 100.

D2SP3SF και 2SP3SF αλλά οι λύσεις τους είναι πολύ χειρότερες όσον αφορά στην αποδοτικότητα ως προς την ενέργεια (ειδικά όταν  $\alpha = 2$ ). Ο αλγόριθμος DISF είναι κάπως καλύτερος όταν  $\alpha = 2$  και μάλλον κακός όταν  $\alpha = 4$ . Ο χρόνος εκτέλεσής τους είναι πολύ μεγάλος. Οι αλγόριθμοι DFF και DFF-2 είναι ακόμα πιο αργοί αλλά φαίνεται να είναι οι καλύτεροι από τους αλγόριθμους που βασίζονται σε εύρεση NWST όσον αφορά στην αποδοτικότητα των λύσεων που επιστρέφουν ως προς την ενέργεια. Συνολικά, όλοι οι αλγόριθμοι που βασίζονται σε εύρεση NWST είναι χειρότεροι από αλγόριθμους που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας. Οι αλγόριθμοι DISF, DFF και DFF-2 περιέχουν κάποιες από τις ιδιότητες αυτές που κάνουν βασικούς αλγόριθμους που χρησιμοποιούν την ιδέα ενδεχόμενης εξοικονόμησης



ενέργειας να έχουν καλύτερη απόδοση, δηλ., επαυξάνουν ένα δένδρο μετάδοσης που περιέχει τη ρίζα συμπεριλαμβάνοντας σε αυτό νέα τερματικά σε κάθε φάση. Δυστυχώς, ο τοπικός στόχος που χρησιμοποιείται σε αλγόριθμους βασισμένους στην εύρεση NWST περιέχει το βάρος αραχνών (spiders) και πηρουνιών (forks) που μπορεί να μην είναι πάντα ανάλογο με την ενέργεια. Επιπλέον, δε βλέπουμε προφανή τρόπο να συμπεριλάβουμε την ιδέα ενδεχόμενης εξοικονόμησης ενέργειας (ή κάποια παραπλήσια) σε αλγόριθμους βασισμένους σε εύρεση NWST. Υπενθυμίζουμε ότι αλγόριθμοι βασισμένοι σε εύρεση NWST αποτελούν παραλλαγές του αλγόριθμου DSF που σχεδιάστηκε για να προσεγγίσει αποδοτικά βέλτιστες λύσεις για το πρόβλημα MEMT στη γενικότερη συμμετρική περίπτωση. Η απόδοση αλγορίθμων βασισμένων σε εύρεση NWST υποδεικνύει ότι η συγκεκριμένη γεωμετρική εκδοχή του προβλήματος που μελετάμε διαθέτει συγκεκριμένες ιδιότητες που αξιοποιούνται καλύτερα από αλγόριθμους με απλούς και διαισθητικούς τοπικούς στόχους.

| $\alpha = 2$ | 10   | 20    | 30    | 40    |
|--------------|------|-------|-------|-------|
| MIP          | 7.16 | 9.61  | 10.99 | 11.97 |
| SP3SF        | 6.94 | 9.09  | 10.28 | 11.18 |
| DSP3SF       | 6.83 | 8.87  | 9.97  | 11.04 |
| DSF          | 8.49 | 11.29 | 12.31 | 12.57 |
| DSF-2        | 8.11 | 10.54 | 11.62 | 12.57 |
| DSF-3        | 8.12 | 10.40 | 11.28 | 11.95 |
| DISF         | 7.72 | 10.06 | 11.00 | 11.79 |
| DFF          | 7.20 | 9.63  | 10.69 | 11.79 |
| DFF-2        | 7.19 | 9.54  | 10.66 | 11.72 |
| iDISF        | 6.94 | 9.15  | 10.18 | 10.95 |
| iDFF         | 7.06 | 9.47  | 10.50 | 11.55 |
| iDFF-2       | 7.04 | 9.39  | 10.44 | 11.53 |

| $\alpha = 4$    | 10   | 20   | 30   | 40    |
|-----------------|------|------|------|-------|
| MIP             | 5.66 | 8.22 | 9.25 | 10.00 |
| SP3SF           | 5.55 | 7.96 | 8.98 | 9.66  |
| DSP3SF          | 5.51 | 7.87 | 9.06 | 9.70  |
| DSF             | 6.52 | 9.13 | 9.99 | 10.27 |
| $\Delta\Phi$ -2 | 5.83 | 8.36 | 9.50 | 10.27 |
| DSF-3           | 6.15 | 8.91 | 9.99 | 10.65 |
| DISF            | 6.01 | 8.76 | 9.74 | 10.49 |
| DFF             | 5.74 | 8.37 | 9.38 | 10.11 |
| DFF-2           | 5.71 | 8.33 | 9.37 | 10.11 |
| iDISF           | 5.57 | 8.05 | 9.11 | 9.79  |
| iDFF            | 5.67 | 8.20 | 9.33 | 10.05 |
| iDFF-2          | 5.65 | 8.20 | 9.32 | 10.00 |

Πίνακας 5.2: Σύγκριση αλγορίθμων βασισμένων στο πρόβλημα NWST και επαναληπτικών τους εκδοχών με τους αλγόριθμους MIP, SP3SF και DSP3SF σε τυχαία στιγμιότυπα με 40 κόμβους,  $\alpha = 2$  (αριστερά) και  $\alpha = 4$  (δεξιά) και μεγέθη ομάδων 10, 20, 30 και 40.

Μελετάμε επίσης επαναληπτικές εκδοχές των αλγορίθμων μας. Από τον ορισμό των αλγορίθμων αυτών, είναι σαφές ότι πάντα συμπεριφέρονται καλύτερα από ότι οι μη περιορισμένες παραλλαγές τους όσον αφορά στην αποδοτικότητα ως προς την ενέργεια, αλλά με κόστος πολλαπλασιαστικό ως προς το μέγεθος των ομάδων στο χρόνο εκτέλεσης. Στην πράξη, στις περισσότερες υλοποιήσεις μας, ο χρόνος εκτέλεσης επαναληπτικών αλγορίθμων πλησιάζει το χρόνο εκτέλεσης των μη περιορισμένων εκδοχών τους (δείτε τον Πίνακα 5.5). Ο αλγόριθμος iDSP3SF υπολογίζει τις πιο αποδοτικές λύσεις ως προς την ενέργεια. Ο αλγόριθμος iD2SP3SF είναι ελαφρώς χειρότερος iDSP3SF ως

προς την αποδοτικότητα ως προς την ενέργεια αλλά ο χρόνος εκτέλεσής του είναι εξαιρετικά μεγάλος. Ο μόνος επαναληπτικός αλγόριθμος με χρόνο εκτέλεσης συγκρίσιμο με αυτόν του αλγορίθμου iDSP3SF είναι ο iSP3SF ο οποίος ουσιαστικά δε βελτιώνει τη μη περιορισμένη παραλλαγή του SP3SF. Συνολικά, οι λύσεις που επιτυγχάνονται από τον αλγόριθμο iDSP3SF είναι μέχρι και 5% (και περίπου 2%) καλύτερες από αυτές του αλγορίθμου iSP3SF όταν  $\alpha = 2$  (και  $\alpha = 4$ , αντίστοιχα). Τα αντίστοιχα αποτελέσματα παραθέτονται στους Πίνακες 5.1. Οι λύσεις που επιτυγχάνονται από επαναληπτικές ακδοχές αλγορίθμων βασισμένων σε εύρεση δένδρων Steiner με βάρη στις κορυφές (NWST) είναι συνήθως πολύ χειρότερες ενώ ο χρόνος εκτέλεσής τους είναι πάρα πολύ μεγάλος. Είναι ενδιαφέρον, πάντως, ότι οι λύσεις που επιτυγχάνονται από τον αλγόριθμο iDISF βελτιώνουν σημαντικά τα αποτελέσματα που επιτυγχάνονται από τον αλγόριθμο DISF. Δυστυχώς, αυτός φαίνεται να είναι ο πιο αργός από όλους τους αλγόριθμους που υλοποιήσαμε.

Η διερεύνηση που παρουσιάζουμε στη συνέχεια απαντά πιθανώς στην ερώτηση γιατί ο αλγόριθμος iDSP3SF είναι καλύτερος από τον iSP3SF ενώ οι αλγόριθμοι DSP3SF και SP3SF υπολογίζουν λύσεις συγκρίσιμης αποδοτικότητας ως προς την ενέργεια. Στον Πίνακα 5.3 παρουσιάζουμε την απόδοση περιορισμένων εκδοχών των αλγορίθμων DSP3SF και SP3SF (παρουσιάζουμε επίσης επιπλέον αποτελέσματα για την περιορισμένη εκδοχή του αλγορίθμου DSPF). Είναι σαφές ότι η περιορισμένη εκδοχή του αλγορίθμου SP3SF υπολογίζει λύσεις με την ίδια περίπου ενέργεια ανεξάρτητα από την παράμετρο περιορισμού επαύξησης, ενώ αυτό δε συμβαίνει για τον αλγόριθμο DSP3SF. Αυτό υποδεικνύει ότι, δεδομένου ενός στιγμιοτύπου του προβλήματος, πολλές διαφορετικές τιμές της παραμέτρου περιορισμού επαύξησης μπορούν να δώσουν τη βέλτιστη λύση ως προς την αποδοτικότητα ως προς την ενέργεια του αλγορίθμου iDSP3SF, ενώ η λύση που επιτυγχάνεται από τον αλγόριθμο iSP3SF είναι οριακά καλύτερη από αυτήν που επιτυγχάνεται από την περιορισμένη εκδοχή του αλγορίθμου SP3SF με τιμή της παραμέτρου περιορισμού επαύξησης ίση με 1.

Η επίδραση αλγορίθμων τοπικής αναζήτησης σε λύσεις που επιτυγχάνονται από επαυξητικούς αλγόριθμους που υλοποιήσαμε είναι σημαντική στην περίπτωση που  $\alpha = 2$  ενώ φαίνεται να είναι οριακή στην περίπτωση που  $\alpha = 4$  (δείτε τον Πίνακα 5.4). Τέτοιοι αλγόριθμοι δεν προσθέτουν σημαντική επιβάρυνση στο συνολικό χρόνο εκτέλεσης και συνήθως οδηγούν σε καλύτερες λύσεις. Ο αλγόριθμος EWMA φαίνεται να είναι κατάλληλος όταν  $\alpha = 2$  και συγκεκριμένα για στιγμιότυπα διάδοσης για τα οποία σχεδιάστηκε αρχικά, ενώ ο αλγόριθμος Sweep φαίνεται να είναι ελαφρώς καλύτερος όταν  $\alpha = 4$ .

|                       |       |       |       |       |       |       |       |       |       |       |       |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\alpha = 2, g = 50$  | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | > 10  |
| DSPF                  | 9.46  | 9.36  | 9.32  | 9.27  | 9.26  | 9.25  | 9.24  | 9.23  | 9.24  | 9.23  | 9.22  |
| SP3SF                 | 9.04  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  | 9.03  |
| DSP3SF                | 9.04  | 9.02  | 8.91  | 8.88  | 8.85  | 8.85  | 8.86  | 8.86  | 8.88  | 8.88  | 8.87  |
| $\alpha = 2, g = 100$ | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | > 10  |
| DSPF                  | 11.61 | 11.43 | 11.32 | 11.25 | 11.21 | 11.18 | 11.15 | 11.15 | 11.12 | 11.11 | 11.07 |
| SP3SF                 | 10.83 | 10.80 | 10.80 | 10.79 | 10.79 | 10.79 | 10.79 | 10.79 | 10.79 | 10.79 | 10.79 |
| DSP3SF                | 10.83 | 10.85 | 10.74 | 10.71 | 10.70 | 10.64 | 10.68 | 10.64 | 10.64 | 10.67 | 10.62 |
| $\alpha = 4, g = 50$  | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | > 10  |
| DSPF                  | 2.82  | 2.83  | 2.82  | 2.82  | 2.82  | 2.82  | 2.82  | 2.82  | 2.82  | 2.82  | 2.82  |
| SP3SF                 | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  | 2.74  |
| DSP3SF                | 2.74  | 2.75  | 2.74  | 2.75  | 2.75  | 2.75  | 2.76  | 2.76  | 2.76  | 2.75  | 2.75  |
| $\alpha = 4, g = 100$ | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | > 10  |
| DSPF                  | 3.52  | 3.52  | 3.52  | 3.53  | 3.53  | 3.52  | 3.52  | 3.52  | 3.52  | 3.52  | 3.51  |
| SP3SF                 | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  | 3.41  |
| DSP3SF                | 3.41  | 3.43  | 3.43  | 3.42  | 3.43  | 3.42  | 3.43  | 3.43  | 3.42  | 3.43  | 3.42  |

Πίνακας 5.3: Η αποδοτικότητα ως προς την ενέργεια περιορισμένων εκδοχών των αλγορίθμων DSPF, SP3SF και DSP3SF σε τυχαία στιγμιότυπα με 100 κόμβους,  $\alpha = 2$  και  $\alpha = 4$ , και μεγέθη ομάδων  $g$  ίσα με 50 και 100 για διαφορετικές τιμές της παραμέτρου που εκφράζει περιορισμούς επαύξησης. Η τελευταία στήλη περιέχει την ενέργεια της καλύτερης λύσης για τιμές της παραμέτρου μεγαλύτερες του 10.

Η βελτίωση σε λύσεις επαυξητικών αλγορίθμων που επιτυγχάνονται μετά την εκτέλεση του αλγορίθμου EWMA και/ή του αλγορίθμου Sweep είναι μεγαλύτερη για τους επαυξητικούς αλγόριθμους που είναι χειρότεροι ως προς την αποδοτικότητα ως προς την ενέργεια και μικρότερη για εκείνους τους αλγόριθμους που επιστρέφουν αποδοτικότερες λύσεις. Συνήθως, η επαναληπτική εκτέλεση αλγορίθμων τοπικής αναζήτησης μπορεί να βελτιώσει παραπάνω μια λύση. Η βελτίωση αυτή αρχίζει να μην είναι σημαντική μετά τις πρώτες δύο οι τρεις εκτελέσεις. Μια ενδιαφέρουσα ερώτηση είναι αν η εκτέλεση του αλγορίθμου EWMA και/ή του αλγορίθμου Sweep μετά τη μη περιορισμένη εκδοχή ενός αλγορίθμου είναι καλύτερη από την επαναληπτική του εκδοχή. Παρατηρήσαμε ότι αυτό ισχύει για τους αλγόριθμους SP3SF και DSP3SF (έχουμε ήδη δει ότι οι επαναληπτικές εκδοχές των αλγορίθμων αυτών δε βελτιώνουν σημαντικά την αποδοτικότητα των λύσεων ως προς την ενέργεια), αλλά κάτι τέτοιο δεν είναι σαφές για τους αλγόριθμους DSPF, DSP3SF, D2SPF και D2SP3SF, ακόμα κι αν συγκρίνουμε τους αλγόριθμους αυτούς ακολουθούμενους από αρκετές κλήσεις των αλγορίθμων EWMA και Sweep με τις επαναληπτικές τους εκδοχές. Επίσης, η εκτέλεση του αλγορίθμου EWMA μετά π.χ., τον αλγόριθμο iDSP3SF βελτιώνει περαιτέρω τη λύση. Τα παραπάνω συνεπάγονται ότι οι

αλγόριθμοι τοπικής αναζήτησης δε μπορούν να υποκαταστήσουν τους επαναληπτικούς αλγόριθμους. Όμως, μπορούν να χρησιμοποιηθούν για να βελτιώσουν ελαφρώς την απόδοσή τους όσον αφορά στην αποδοτικότητα ως προς την ενέργεια των λύσεων που επιτυγχάνονται.

Συμπερασματικά, ο αλγόριθμος iDSP3SF ακολουθούμενος από τον αλγόριθμο EWMA φαίνεται να δίνει τις πιο αποδοτικές ως προς την ενέργεια λύσεις. Οι αλγόριθμοι SP3SF και DSP3SF ακολουθούμενοι από αλγόριθμους τοπικής αναζήτησης παρέχουν έναν καλό συμβιβασμό μεταξύ αποδοτικότητας ως προς την ενέργεια και χρόνου εκτέλεσης.

| $\alpha = 2$ | 10   |      |      |      | 50    |      |      |      | 100   |       |       |       |
|--------------|------|------|------|------|-------|------|------|------|-------|-------|-------|-------|
| SPF          | 5.27 | 5.13 | 5.15 | 5.04 | 10.07 | 9.55 | 9.40 | 9.06 | 12.45 | 11.76 | 11.00 | 10.77 |
| MIP          | 4.98 | 4.91 | 4.92 | 4.87 | 9.47  | 9.09 | 9.06 | 8.82 | 11.61 | 10.93 | 10.72 | 10.46 |
| DSPF         | 4.93 | 4.87 | 4.89 | 4.85 | 9.22  | 8.98 | 8.90 | 8.75 | 11.07 | 10.70 | 10.51 | 10.31 |
| iDSPF        | 4.89 | 4.84 | 4.85 | 4.81 | 9.07  | 8.82 | 8.79 | 8.63 | 10.92 | 10.57 | 10.37 | 10.19 |
| DSP3SF       | 4.87 | 4.87 | 4.84 | 4.83 | 8.88  | 8.88 | 8.74 | 8.71 | 10.63 | 10.63 | 10.30 | 10.23 |
| iDSP3SF      | 4.81 | 4.81 | 4.79 | 4.78 | 8.58  | 8.58 | 8.48 | 8.46 | 10.28 | 10.28 | 10.05 | 10.02 |
| SP3SF        | 4.92 | 4.92 | 4.89 | 4.89 | 9.03  | 9.02 | 8.85 | 8.79 | 10.79 | 10.78 | 10.39 | 10.32 |
| iSP3SF       | 4.92 | 4.92 | 4.89 | 4.89 | 9.03  | 9.02 | 8.85 | 8.79 | 10.79 | 10.78 | 10.39 | 10.31 |
| $\alpha = 4$ | 10   |      |      |      | 50    |      |      |      | 100   |       |       |       |
| SPF          | 1.51 | 1.49 | 1.50 | 1.48 | 2.89  | 2.81 | 2.85 | 2.79 | 3.61  | 3.49  | 3.53  | 3.45  |
| MIP          | 1.48 | 1.47 | 1.48 | 1.47 | 2.82  | 2.75 | 2.79 | 2.74 | 3.52  | 3.41  | 3.49  | 3.40  |
| DSPF         | 1.49 | 1.48 | 1.48 | 1.47 | 2.82  | 2.76 | 2.80 | 2.76 | 3.52  | 3.42  | 3.49  | 3.42  |
| iDSPF        | 1.47 | 1.46 | 1.47 | 1.46 | 2.78  | 2.73 | 2.76 | 2.73 | 3.46  | 3.37  | 3.43  | 3.36  |
| DSP3SF       | 1.47 | 1.47 | 1.46 | 1.46 | 2.75  | 2.75 | 2.75 | 2.74 | 3.43  | 3.43  | 3.43  | 3.42  |
| iDSP3SF      | 1.45 | 1.45 | 1.45 | 1.45 | 2.70  | 2.69 | 2.69 | 2.69 | 3.34  | 3.34  | 3.34  | 3.33  |
| SP3SF        | 1.47 | 1.47 | 1.46 | 1.46 | 2.74  | 2.74 | 2.74 | 2.73 | 3.41  | 3.41  | 3.40  | 3.39  |
| iSP3SF       | 1.47 | 1.47 | 1.46 | 1.46 | 2.74  | 2.74 | 2.74 | 2.73 | 3.41  | 3.41  | 3.40  | 3.39  |

Πίνακας 5.4: Επίδραση αλγορίθμων οπτικής αναζήτησης σε επαυξητικούς αλγόριθμους για τυχαία στιγμιότυπα με 100 κόμβους,  $\alpha = 2$  και  $\alpha = 4$ , και μεγέθη ομάδας 10, 50 και 100. Οι τέσσερις στήλες για κάθε μέγεθος ομάδας συμβολίζουν την ενέργεια του αλγορίθμου, τον αλγόριθμο ακολουθούμενο από Sweep, τον αλγόριθμο ακολουθούμενο από EWMA, και τον αλγόριθμο ακολουθούμενο από εκτελέσεις των EWMA, Sweep και EWMA, αντίστοιχα.

| Αλγόριθμος | Χρόνος   | Αλγόριθμος | Χρόνος   | Αλγόριθμος | Χρόνος    |
|------------|----------|------------|----------|------------|-----------|
| SPF        | 12 msec  | iDSPF      | 95 msec  | DSF-3      | 1.2 sec   |
| MIP        | 14 msec  | iD2SPF     | 6 sec    | DISF       | 11.6 sec  |
| DSPF       | 6 msec   | iSP3SF     | 250 msec | DFF        | 15.7 sec  |
| D2SPF      | 265 msec | iDSP3SF    | 130 msec | DFF-2      | 18.1 sec  |
| SP3SF      | 22 msec  | i2SP3SF    | 10.7 sec | iDISF      | 349 sec   |
| DSP3SF     | 6 msec   | iD2SP3SF   | 6.7 sec  | iDFF       | 137.6 sec |
| 2SP3SF     | 1 sec    | DSF        | 1.2 sec  | iDFF-2     | 194.8 sec |
| D2SP3SF    | 280 msec | DSF-2      | 1.2 sec  |            |           |

Πίνακας 5.5: Χρόνος εκτέλεσης των αλγορίθμων σε τυχαία στιγμιότυπα με 40 κόμβους και μέγεθος ομάδας 40.



# Κεφάλαιο 6

## Επίλογος

Στην παρούσα διδακτορική διατριβή, μελετήσαμε προβλήματα επικοινωνίας που ανακύπτουν σε ασύρματα αδόμητα δίκτυα, δηλ. δίκτυα που δε βασίζονται σε ενσύρματη τηλεπικοινωνιακή υποδομή αλλά στην ύπαρξη ασύρματων κόμβων (εκπομπών και δεκτών) με περιορισμένα αποθέματα ενέργειας και στην επικοινωνία μεταξύ τους. Σε τέτοια δίκτυα, η ενέργεια αποτελεί έναν πολύ κρίσιμο πόρο και η αποδοτική της διαχείριση είναι εξαιρετικά σημαντική ακόμα και για την ίδια την ύπαρξη του δικτύου. Παρουσιάσαμε αποδοτικές τεχνικές για την επίλυση προβλημάτων μετάδοσης (multicasting) με τρόπο αποδοτικό ως προς την κατανάλωση ενέργειας καθώς επίσης και τεχνικές για ενεργειακά αποδοτική ομαδική επικοινωνία και διατήρηση συνεκτικότητας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης.

Μελετήσαμε το πρόβλημα κάλυψης συνόλου και ενδιαφέρουσες παραλλαγές του που σχετίζονται άμεσα με το πρόβλημα της ενεργειακά αποδοτικής μετάδοσης (multicasting) σε ασύρματα δίκτυα. Παρουσιάσαμε νέους συνδυαστικούς προσεγγιστικούς αλγόριθμους για το πρόβλημα κάλυψης με σύνολα μεγέθους το πολύ  $k$ . Οι αλγόριθμοί μας επεκτείνουν περαιτέρω προηγούμενες προσεγγίσεις - που βασίζονταν σε επεκτάσεις του άπληστου αλγόριθμου μέσω αποδοτικού χειρισμού μικρών συνόλων - χρησιμοποιώντας την ιδέα του υπολογισμού μεγάλων ομάδων στοιχείων και στη συνέχεια της ομαδοποίησής τους σε σύνολα μεγάλου μεγέθους. Τα αποτελέσματα που επιτύχαμε βελτιώνουν τα καλύτερα γνωστά φράγματα προσέγγισης για το πρόβλημα κάλυψης με σύνολα μεγέθους το πολύ  $k$  για κάθε τιμή του  $k \geq 6$ . Η τεχνική που χρησιμοποιήσαμε για την ανάλυση παρουσιάζει ανεξάρτητο ενδιαφέρον: το πάνω φράγμα για τον παράγοντα προσέγγισης επιτυγχάνεται φράσσοντας την

αντικειμενική τιμή ενός γραμμικού προγράμματος η οποία ‘αποκαλύπτει’ το λόγο προσέγγισης του υπό εξέταση αλγορίθμου (*factor-revealing linear program*). Οι αλγόριθμοί μας καθώς και οι προηγούμενες γνωστές βελτιώσεις του άπληστου αλγορίθμου επιτυγχάνουν λόγο προσέγγισης ίσο με  $H_k$  μείον μια μικρή σταθερά. Το γνωστό κάτω φράγμα για το λόγο προσέγγισης είναι  $H_k - O(\log \log k)$  [76]. Αποτελεί ενδιαφέρον θεωρητικό πρόβλημα να κλείσει το χάσμα σχετικά με τους όρους δεύτερης τάξης. Σημειώνουμε ότι η τεχνική ανάλυσης συνδυαστικών αλγορίθμων με χρήση γραμμικών προγραμμάτων παρουσιάζει γενικότερο ενδιαφέρον και θα μπορούσε να εφαρμοστεί και σε άλλα προβλήματα βελτιστοποίησης.

Παρουσιάσαμε έναν απλό αλγόριθμο για το πρόβλημα εύρεσης μέγιστου γεννητικού δάσους αστέρων. Λαμβάνοντας υπόψη το γεγονός ότι το πρόβλημα αποτελεί ειδική περίπτωση του συμπληρωματικού προβλήματος κάλυψης συνόλου, προσαρμόσαμε έναν αλγόριθμο των Duh και Furer για την επίλυσή του. Αποδείξαμε ότι ο αλγόριθμός μας υπολογίζει  $193/240 \approx 0.804$ -προσεγγιστικά γεννητικά δάση αστέρων. Το αποτέλεσμα αυτό βελτιώνει ένα προηγούμενο κάτω φράγμα με τιμή 0.71 των Chen και άλλων. Αν και ο αλγόριθμος είναι καθαρά συνδυαστικός, η ανάλυσή μας ορίζει ένα γραμμικό πρόγραμμα που χρησιμοποιεί μια παράμετρο  $f$  το οποίο είναι εφικτό για τιμές της παραμέτρου  $f$  που δεν είναι μικρότερες από το λόγο προσέγγισης του αλγορίθμου. Η ανάλυση είναι αυστηρή και, το ενδιαφέρον είναι, ότι μπορεί να εφαρμοστεί και σε συμπληρωματικές εκδοχές του προβλήματος κάλυψης συνόλου όπως η εξοικονόμηση χρωμάτων. Δίνει την ίδια εγγύηση προσέγγισης με τιμή  $193/240$  που οριακά βελτιώνει το προηγούμενο γνωστό κάτω φράγμα των Duh και Furer. Η περαιτέρω βελτίωση αυτού του φράγματος αποτελεί σημαντικό ανοιχτό πρόβλημα. Αποδείξαμε επίσης ότι, γενικά, μια κλάση αλγορίθμων τοπικής αναζήτησης δε δίνουν καλύτερα από  $1/2$ -προσεγγιστικά γεννητικά δάση αστέρων.

Μελετήσαμε προβλήματα επικοινωνίας σε ασύρματα δίκτυα που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης. Σε τέτοια δίκτυα, η ενεργοποίηση ενός μέσου ασύρματης διασύνδεσης επιφέρει ένα κόστος που αντανάκλα την ενέργεια που καταναλώνεται όταν κάποιος κόμβος χρησιμοποιεί το μέσο αυτό. Διακρίναμε μεταξύ της συμμετρικής και της μη συμμετρικής περίπτωσης, με βάση τον το κόστος ενεργοποίησης για κάθε ασύρματο μέσο διασύνδεσης είναι το ίδιο για όλους τους κόμβους ή όχι. Για τη συμμετρική περίπτωση, παρουσιάσαμε έναν  $(3/2 + \epsilon)$ -προσεγγιστικό αλγόριθμο για το πρόβλημα πλήρους διασύνδεσης με ελάχιστο κόστος ενεργοποίησης, βελτιώνοντας ένα προηγούμενο φράγμα με τιμή 2. Για τη μη συμμετρική



περίπτωση, αποδείξαμε ότι το πρόβλημα διασύνδεσης δεν είναι προσεγγίσιμο στα πλαίσια ενός παράγοντα υπολογαριθμικού ως προς το πλήθος των κόμβων και παρουσιάσαμε ένα λογαριθμικό προσεγγιστικό αλγόριθμο για μια γενικότερη περίπτωση που μοντελοποιεί την ομαδική επικοινωνία. Τα ανοιχτά προβλήματα στην περιοχή των ασύρματων δικτύων που υποστηρίζουν πολλαπλά μέσα ασύρματης διασύνδεσης είναι πολλά. Ειδικότερα, για τη συμμετρική περίπτωση, είναι ενδιαφέρον να βελτιωθούν τα άνω φράγματα που παρουσιάσαμε. Επιπλέον, ο αλγόριθμος για το πρόβλημα της πλήρους διασύνδεσης είναι ιδιαίτερα πολύπλοκος. Θα ήταν ενδιαφέρον να αναζητηθούν απλούστεροι αλγόριθμοι με την ίδια ή καλύτερη απόδοση. Επίσης, η θεώρηση διαφορετικών ιδιοτήτων συνεκτικότητας αποκαλύπτει μια σειρά από ενδιαφέροντα προβλήματα συνδυαστικής βελτιστοποίησης που συνδέονται με την ελαχιστοποίηση της κατανάλωσης ενέργειας σε αυτά τα δίκτυα.

Σχετικά με το πρόβλημα αποδοτικής ως προς την κατανάλωση ενέργειας μετάδοσης (multicasting) σε ασύρματα αδόμητα δίκτυα, μελετήσαμε δύο κατηγορίες αλγορίθμων της βιβλιογραφίας: *αλγόριθμους τοπικής αναζήτησης (local search algorithms)* και *επαυξητικούς αλγόριθμους (augmentation algorithms)*. Παρουσιάσαμε τα γενικά χαρακτηριστικά και των δύο κατηγοριών και υλοποιήσεις πολλών αλγορίθμων, υπαρχόντων και νέων. Οι υλοποιήσεις μας περιλαμβάνουν αλγόριθμους που προσεγγίζουν το πρόβλημα εύρεσης δένδρου διάδοσης ελάχιστης ενέργειας MEMT στη γενικότερη συμμετρική περίπτωση καθώς και αλγόριθμους που είναι πιο διαισθητικοί για τη γεωμετρική περίπτωση. Συγκρίναμε πειραματικά τους αλγόριθμους αυτούς όσον αφορά στην αποδοτικότητα ως προς την ενέργεια των λύσεων που επιτυγχάνουν για τυχαία στιγμιότυπα του προβλήματος MEMT στο Ευκλείδειο επίπεδο και παρουσιάσαμε αξιολόγηση των χρόνων εκτέλεσης των αλγορίθμων. Τα αποτελέσματά μας δείχνουν ότι κάποιοι από τους νέους αλγόριθμους έχουν πολύ καλή συμπεριφορά όταν εφαρμόζονται σε τυχαία στιγμιότυπα. Δυστυχώς, η πολυπλοκότητα της λειτουργίας των αλγορίθμων αυτών καθιστά ιδιαίτερα δύσκολη τη θεωρητική τους ανάλυση. Επομένως, αποτελεί σημαντικό ανοιχτό πρόβλημα ο σχεδιασμός αλγορίθμων με αποδεδειγμένα καλό λόγο απόδοσης σε συνθήκες χειρότερης περίπτωσης.



# Βιβλιογραφία

- [1] A. Agra, D. Cardoso, O. Cerfeira, and E. Rocha. A spanning star forest model for the diversity problem in automobile industry. *In Proceedings of ECCO XVII*, 2005.
- [2] A. Agrawal, P. Klein, and R. Ravi. When trees collide: an approximation algorithm for generalized Steiner tree Problem on Networks. *In Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC 91)*, pp. 131-144, 1991.
- [3] E. Althaus, G. Călinescu, I. Măndoiu, S. Prasad, N. Tchervenski and A. Zelikovsky. Power Efficient Range Assignment in Ad-Hoc Wireless Networks. *In Proc. of the IEEE Wireless Communications and Networking Conference (WCNC '03)*, IEEE Computer Society Press, pp. 1889-1894, 2003.
- [4] S. Athanassopoulos, I. Caragiannis, and C. Kaklamanis. Analysis of approximation algorithms for  $k$ -set cover using factor-revealing linear programs. *In Proc. of the 16th International Symposium on Fundamentals of Computation Theory (FCT '07)*, LNCS 4639, Springer, pp. 52-63, 2007. *Theory of Computing Systems*, 2009, to appear.
- [5] S. Athanassopoulos, I. Caragiannis, C. Kaklamanis, and P. Kanellopoulos. Experimental Comparison of Algorithms for Energy-Efficient Multicasting in Ad Hoc Networks. *In Proc. of the 3rd International Conference for Ad Hoc Networks and Wireless (ADHOC-NOW '04)*, LNCS 3158, Springer, pp. 183-196, 2004.
- [6] S. Athanassopoulos, I. Caragiannis, C. Kaklamanis, M. Kyropoulou. An improved approximation bound for spanning star forest and color saving. *In Proc. of the 34th International Symposium on Mathematical*

*Foundations of Computer Science (MFCS '09)*, LNCS, Springer, 2009, to appear.

- [7] S. Athanassopoulos, I. Caragiannis, C. Kaklamanis, E. Papaioannou. Energy-efficient communication in multi-interface wireless networks. In *Proc. of the 34th International Symposium on Mathematical Foundations of Computer Science (MFCS '09)*, LNCS, Springer, 2009, to appear.
- [8] G. Ausiello, A. D'Atri, and M. Protasi. Structure preserving reductions among convex optimization problems. *Journal of Computer and System Sciences*, 21, pp. 136-153, 1980.
- [9] G. Ausiello, A. Marchetti-Spaccamela, and M. Protasi. Toward a unified approach for the classification of NP-complete optimization problems. *Theoretical Computer Science*, 12, pp. 83-96, 1980.
- [10] P. Bahl, A. Adya, J. Padhye, and A. Walman. Reconsidering wireless systems with multiple radios. *SIGCOMM Comput. Commun. Rev.*, 34(5), pp. 39-46, 2004.
- [11] V. Berry, S. Guillemot, F. Nicholas, and C. Paul. On the approximation of computing evolutionary trees. In *Proceedings of the 11th Annual International Computing and Combinatorics Conference (COCOON '05)*, LNCS, Springer, pp. 115-123, 2005.
- [12] M. Cagalj, J.P. Hubaux and C. Enz. Minimum-Energy Broadcast in All-Wireless Networks: NP-completeness and Distribution Issues. In *Proc. of the 8th ACM International Conference on Mobile Networking and Computing (Mobicom '02)*, pp. 172-182, 2002.
- [13] G. Călinescu, S. Kapoor, A. Olshevsky and A. Zelikovsky. Network Lifetime and Power Assignment in Ad-Hoc Wireless Networks. In *Proc. of the 11th Annual European Symposium on Algorithms (ESA '03)*, LNCS 2832, Springer, pp. 114-126, 2003.
- [14] G. Călinescu, I. Măndoiu and A. Zelikovsky. Symmetric Connectivity with Minimum Power Consumption in Radio Networks. In *Proc. of the 2nd IFIP International Conference on Theoretical Computer Science*, pp. 119-130, 2002.

- [15] I. Caragiannis. Wavelength management in WDM rings to maximize the number of connections. In *Proceedings of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS '07)*, LNCS 4393, Springer, pp. 61-72, 2007.
- [16] I. Caragiannis, C. Kaklamanis and P. Kanellopoulos. New Results for Energy-Efficient Broadcasting in Wireless Networks. In *Proc. of the 13th Annual International Symposium on Algorithms and Computation (ISAAC '02)*, pp. 332-343, 2002.
- [17] I. Caragiannis, C. Kaklamanis, and P. Kanellopoulos. Energy-efficient wireless network design. *Theory of Computing Systems*, 39(5), pp. 593-617, 2006.
- [18] M. Cardei and D.-Z. Du. Improving Wireless Sensor Network Lifetime through Power Aware Organization. *Wireless Networks*, Volume 11 , Issue 3, pp. 333 - 340, 2005.
- [19] J. Cartigny, D. Simplot, I. Stojmenovic. Localized minimum energy broadcasting in ad hoc networks. In *Proc. of IEEE INFOCOM 2003*, 2003.
- [20] D. Cavalcanti, H. Gossain, and D. Agrawal. Connectivity in multi-radio, multi-channel heterogeneous ad hoc networks. In *Proceedings of the IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 05)*, IEEE, pp. 1322-1326, 2005.
- [21] D. Chakrabarty and G. Goel. On the approximability of budgeted allocations and improved lower bounds for submodular welfare maximization and GAP. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08)*, pp. 687-696, 2008.
- [22] M. Charikar, C. Chekuri, T.-Y. Cheung, Z. Dai, A. Goel, S. Guha and M. Li. Approximation Algorithms for Directed Steiner Problems. *Journal of Algorithms*, 33(1):73-91, 1999.
- [23] N. Chen, R. Engelberg, C. T. Nguyen, P. Raghavendra, A. Rudra, and G. Singh. Improved approximation algorithms for the spanning star forest problem. In *Proceedings of APPROX-RANDOM 2007*, LNCS 4627, Springer, pp. 44-58, 2007.

- [24] V. Chvátal. A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, 4, pp. 233-235, 1979.
- [25] A. E. F. Clementi, P. Crescenzi, P. Penna, G. Rossi and P. Vocca. On the Complexity of Computing Minimum Energy Consumption Broadcast Subgraphs. In *Proc. of the 18th Annual Symposium on Theoretical Aspects of Computer Science (STACS '01)*, LNCS 2010, Springer, pp. 121-131, 2001.
- [26] A. E. F. Clementi, M. Di Ianni and R. Silvestri. The Minimum Broadcast Range Assignment Problem on Linear Multi-Hop Wireless Networks. *Theoretical Computer Science*, 1-3(299): 751-761, 2003.
- [27] A. E. F. Clementi, P. Penna, A. Ferreira, S. Perennes and R. Silvestri. The Minimum Range Assignment Problem on Linear Radio Networks. *Algorithmica*, 35(2): 95-110, 2003.
- [28] A. E. F. Clementi, P. Penna and R. Silvestri. Hardness Results for the Power Range Assignment Problem in Packet Radio Networks. In *Proc. of Randomization, Approximation, and Combinatorial Optimization (RANDOM/APPROX '99)*, LNCS 1671, Springer, pp. 197-208, 1999.
- [29] T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein. Introduction to Algorithms. *The MIT Press*, Second Edition, 2001.
- [30] M. Demange, P. Grisoni, and V. T. Paschos. Differential approximation algorithms for some combinatorial optimization problems. *Theoretical Computer Science*, 209, pp. 107-122, 1998.
- [31] M. Demange and V. T. Paschos. On an approximation measure founded on the links between optimization and polynomial approximation theory. *Theoretical Computer Science*, 158, pp. 117-141, 1996.
- [32] R. Draves, J. Padhye, and B. Zill. Routing in multi-radio, multi-hop wireless mesh networks. In *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom 04)*, ACM, pp. 114-128, 2004.
- [33] S. E. Dreyfus and R. A. Wagner. The Steiner problem in graphs. *Networks*, vol. 1 (1971), pp. 195-207.

- [34] R. Duh and M. Fürer. Approximation of  $k$ -set cover by semi local optimization. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC '97)*, pp. 256-264, 1997.
- [35] J. Edmonds. Optimum Branchings. *Journal of Research of the National Bureau of Standards*, 71B:233-240, 1967.
- [36] P. Erdős and H. Sachs. Reguläre Graphen gegebener Tailienweite mit minimaler Knotenzahl, *Wiss. Z. Univ. Halle, Math.-Nat.*, 12, pp. 251-258, 1963.
- [37] A. Faragó and S. Basagni. The effect of multi-radio nodes on network connectivity: a graph theoretic analysis. In *Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 08)*, IEEE, 2008.
- [38] U. Feige. A Threshold of  $\ln n$  for Approximating Set Cover. *Journal of the ACM*, Vol. 45(4), pp. 634-652, 1998.
- [39] M. X. Goemans and D. P. Williamson. A general approximation technique for constrained forest problems. *SIAM Journal on Computing*, 24, pp. 296-317, 1995.
- [40] O. Goldschmidt, D. Hochbaum, and G. Yu. A modified greedy heuristic for the set covering problem with improved worst case bound. *Information Processing Letters*, 48, pp. 305-310, 1993.
- [41] S. Guha and S. Khuller. Improved Methods for Approximating Node Weighted Steiner Trees and Connected Dominating Sets. *Information and Computation*, 150(1), pp. 57-74, 1999.
- [42] M. Hajiaghayi, N. Immorlica and V. Mirrokni. Power Optimization in Fault-Tolerant Topology Control Algorithms for Wireless Multi-hop Networks. In *Proc. of the 9th Annual International Conference on Mobile Computing and Networking (MOBICOM '03)*, 2003.
- [43] M. M. Halldórsson. Approximating discrete collections via local improvements. In *Proceedings of the 6th Annual ACM/SIAM Symposium on Discrete Algorithms (SODA '95)*, pp. 160-169, 1995.

- [44] M. M. Halldórsson. Approximating  $k$ -set cover and complementary graph coloring. In *Proceedings of the 5th Conference on Integer Programming and Combinatorial Optimization (IPCO '96)*, LNCS 1084, Springer, pp. 118-131, 1996.
- [45] E. Halperin and R. Krauthgamer. Polylogarithmic Inapproximability. In *Proc. of the 35th Annual ACM Symposium on Theory of Computing (STOC '03)*, pp. 585-594, 2003.
- [46] R. Hassin and S. Khuller.  $z$ -Approximations. *Journal of Algorithms*, 41(2), pp. 429-442, 2001.
- [47] E. Hazan, S. Safra, and O. Schwartz. On the complexity of approximating  $k$ -set packing. *Computational Complexity*, 15(1), pp. 20-39, 2006.
- [48] C. A. J. Hurkens and A. Schrijver. On the size of systems of sets every  $t$  of which have an SDR, with an application to the worst-case ratio of heuristics for packing problems. *SIAM Journal on Discrete Mathematics*, 2(1), pp. 68-72, 1989.
- [49] F. K. Hwang and D. S. Richards. Steiner tree problems. *Networks*, Vol. 22, No. 1, pp. 55-90 (1992).
- [50] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed Diffusion: A Scalable and Robust communication Paradigm for Sensor Networks. In *Proc. of the 6th annual ACM International Conference on Mobile Computing and Networking (MobiCom 00)*, pp. 56-67, 2000.
- [51] K. Jain, M. Mahdian, E. Markakis, A. Saberi, and V. V. Vazirani. Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. *Journal of the ACM*, 50(6), pp. 795-824, 2003.
- [52] D. S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9, pp. 256-278, 1974.
- [53] V. Kann. Maximum bounded 3-dimensional matching is MAX SNP-complete. *Information Processing Letters*, 37, pp. 27-35, 1991.
- [54] L. M. Kirousis, E. Kranakis, D. Krizanc and A. Pelc. Power Consumption in Packet Radio Networks. *Theoretical Computer Science*, 243(1-2):289-305, 2000.



- [55] R. Klasing, A. Kosowski, and A. Navarra. Cost minimisation in wireless networks with bounded and unbounded number of interfaces. *Networks*, 53(3), pp. 266-275, 2009.
- [56] P.N. Klein and R. Ravi. A Nearly Best Possible Approximation Algorithm for Node-Weighted Steiner Trees. *Journal of Algorithms*, 19(1), pp. 104-115, 1995.
- [57] A. Kosowski, A. Navarra, and M. C. Pinotti. Connectivity in multi-interface networks. In *Proceedings of the 4th International Symposium on Trustworthy Global Computing (TGC 08)*, LNCS 5474, Springer, pp. 159-173, 2009.
- [58] L. Kou, G. Markowsky and L. Berman. A fast algorithm for Steiner trees. *Acta Informatica*, vol 15 (1981), pp. 141-145.
- [59] A. Levin. Approximating the unweighted  $k$ -set cover problem: greedy meets local search. *SIAM Journal on Discrete Mathematics*, 23(1), pp. 251-264, 2009.
- [60] Y. Li. Energy management in wireless networks. PhD thesis. University of Minnesota, 2005.
- [61] W. Liang. Constructing Minimum-Energy Broadcast Trees in Wireless Ad Hoc Networks. In *Proc. of 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '02)*, pp. 112-122, 2002.
- [62] E. Lloyd, R. Liu, M. Marathe, R. Ramanathan and S.S. Ravi. Algorithmic Aspects of Topology Control Problems for Ad hoc Networks. In *Proc. of 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '02)*, pp. 123-134, 2002.
- [63] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13, pp. 383-390, 1975.
- [64] P. Mavinkurve, H.Q. Ngo and H. Mensa. MIP3S: Algorithms for Power-conserving Multicasting in Wireless Ad Hoc Networks. In *Proc. of the 11th IEEE International Conference on Networks (ICON '03)*, 2003.

- [65] K. Melhorn. A faster approximation algorithm for the Steiner tree problem in graphs. *Information Processing Letters* vol. 27(3)(1988), pp. 125-128.
- [66] C. T. Nguyen, J. Shen, M. Hou, L. Sheng, W. Miller, and L. Zhang. Approximating the spanning star forest problem and its application to genomic sequence alignment. *SIAM Journal on Computing*, 38(3), pp. 946-962, 2008.
- [67] J. Plesnik. A bound for the Steiner tree problem in graphs. *Math. Slovaca*, vol. 31 (1981), pp. 155-163.
- [68] H.J. Prömel and A. Steger. A new approximation algorithm for the Steiner tree problem with performance ratio  $5/3$ . *Journal of Algorithms*, 36, pp. 89-101, 2000.
- [69] T. S. Rappaport. *Wireless Communications: Principles and Practices*. Prentice Hall, 1996.
- [70] V. J. Rayward-Smith. The computation of nearly minimal Steiner trees in graphs. *Internat. J. Math. Ed. Sci. Tech.*, 14(1) pp. 15-23 (1983).
- [71] V. J. Rayward-Smith and A. Clare. On finding Steiner vertices. *Networks*, 16, pp. 283-294 (1986).
- [72] R. Raz, S. Safra. A Sub-Constant Error-Probability Low-Degree Test, and a Sub-Constant Error-Probability PCP Characterization of NP. In *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing (STOC 97)*, pp. 475-484, 1997.
- [73] G. Robins and A. Zelikovsky. Improved Steiner Tree Approximations in Graphs. In *Proc. of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '00)*, pp. 770-779, 2000.
- [74] P. Slavik. A tight analysis of the greedy algorithm for set cover. *Journal of Algorithms*, 25, pp. 237-254, 1997.
- [75] H. Takahashi and A. Matsuyama. An approximate solution for the Steiner problem in graphs. *Math. Japonica*, vol.24 (1980), pp. 573-577.

- [76] L. Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC '01)*, pp. 453-461, 2001.
- [77] V. Verma, A. Chandak and H.Q. Ngo. DIP3S: A Distributive Routing Algorithm for Power-Conserving Broadcasting in Wireless Ad Hoc Networks. In *Proc. of the Fifth IFIP-TC6 International Conference on Mobile and Wireless Communications Networks (MWCN '03)*, pp. 159-162, 2003.
- [78] P.-J. Wan, G. Călinescu, X.-Y. Li and O. Frieder. Minimum-Energy Broadcasting in Static Ad Hoc Wireless Networks. *Wireless Networks*, 8(6):607-617, 2002.
- [79] J. E. Wieselthier, G. D. Nguyen and A. Ephremides. On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks. In *Proc. of IEEE INFOCOM 2000*, pp. 585-594, 2000.
- [80] P. Winter. Steiner problem in networks: a survey. *BIT* 25 (1985), pp. 485-496.
- [81] J. Wu and F. Dai. Broadcasting in Ad Hoc Networks Based on Self-Pruning. In *Proc. of the 22nd Annual Joint Conference of IEEE Communication and Computer Society*, (CD-ROM) 2003.
- [82] J. Wu, F. Dai, M. Gao, and I. Stojmenovic. On Calculating Power-Aware Connected Dominating Sets for Efficient Routing in Ad Hoc Wireless Networks. *Journal of Communications and Networks*, vol. 1, pp. 1-12, 2002.
- [83] J. Wu and H. Li. On Calculating Connected Dominating Set for Efficient Routing in Ad Hoc Wireless Networks. In *Proc. of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIALM 99)* pp. 7-14, 1999.
- [84] Y. F. Wu, P. Widmayer and C. K. Wong. A faster approximation algorithm for the Steiner problem in graphs. *Acta Informatica*, vol. 23 (1986), pp. 321-331.
- [85] A. Z. Zelikovsky. The  $11/6$ -approximation algorithm for the Steiner problem on networks. *Institute of Mathematics, Kishinev, USSR*.

- [86] L. Zosin and S. Khuller. On Directed Steiner Trees. In *Proc. of the 13th Annual ACM/SIAM Symposium on Discrete Algorithms (SODA '02)*, pp. 59-63, 2002.

