Θεωρία Γραφημάτων: Σύνολα Ανεξαρτησίας, Σύνολα Κάλυψης, και Χρωματικός Αριθμός

Δημήτρης Φωτάκης

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων
Πανεπιστήμιο Αιγαίου, 83200 Καρλόβασι, Σάμος
Email: fotakis@aegean.gr

1 Βασικοί Ορισμοί

Έστω γράφημα $G(V, E)$. Ένα σύνολο κορυφών $S \subseteq V$ ονομάζεται *σύνολο ανεξαρτησίας* (independent set) αν δεν υπάρχουν ακμές μεταξύ αυτών των κορυφών. Ένα σύνολο ανεξαρτησίας είναι *μέγιστο* (Maximum Independent Set - MIS) όταν δεν υπάρχει άλλο μεγαλύτερο σύνολο ανεξαρτησίας στο γράφημα. Ο αριθμός των κορυφών (ή το μέγεθος) του μεγαλύτερου συνόλου ανεξαρτησίας ονομάζεται *αριθμός ανεξαρτησίας* (independence number) του γραφήματος G και συμβολίζεται με $\alpha(G)$. Ένα σύνολο κορυφών S είναι σύνολο ανεξαρτησίας στο G αν και μόνο αν το S είναι *κλίκα* (clique), δηλαδή πλήρες υπογράφημα, στο συμπληρωματικό γράφημα $\overline{G}$.

Ένα σύνολο κορυφών $C \subseteq V$ ονομάζεται *σύνολο κάλυψης* ή *κάλυμμα κορυφών* (vertex cover) όταν κάθε ακμή του γραφήματος έχει τουλάχιστον ένα από τα άκρα της στο C. Ένα σύνολο κάλυψης είναι *ελάχιστο* (Minimum Vertex Cover - MVC) όταν δεν υπάρχει άλλο μικρότερο σύνολο κάλυψης στο γράφημα. Ο αριθμός των κορυφών (ή το μέγεθος) του μικρότερου συνόλου κάλυψης ονομάζεται *αριθμός κάλυψης* (covering number) του γραφήματος και συμβολίζεται με $\beta(G)$.

Πρόταση 1. *Έστω γράφημα $G(V, E)$. Ένα σύνολο κορυφών $S \subseteq V$ είναι σύνολο ανεξαρτησίας του G αν και μόνο αν το σύνολο $V \setminus S$ είναι σύνολο κάλυψης.*

Απόδειξη. Εξ' ορισμού, το S είναι σύνολο ανεξαρτησίας αν και μόνο αν δεν υπάρχει καμία ακμή που έχει και τα δύο άκρα της στο S. Αυτό συμβαίνει αν και μόνο αν κάθε ακμή έχει τουλάχιστον ένα από τα άκρα της στο $V \setminus S$, δηλαδή αν και μόνο αν το $V \setminus S$ είναι σύνολο κάλυψης. □

Πρόταση 2. *Σε κάθε γράφημα $G(V, E)$, $\alpha(G) + \beta(G) = |V|$.*

Απόδειξη. Έστω S ένα μέγιστο σύνολο ανεξαρτησίας του G. Εξ' ορισμού είναι $|S| = \alpha(G)$. Από την Πρόταση 1, το $V \setminus S$ είναι ένα σύνολο κάλυψης, και επομένως $\beta(G) \leq n - \alpha(G) \Rightarrow \alpha(G) \leq n - \beta(G)$.

Έστω C ένα ελάχιστο σύνολο κάλυψης του G. Εξ' ορισμού είναι $|C| = \beta(G)$. Από την Πρόταση 1, το $V \setminus C$ είναι ένα σύνολο ανεξαρτησίας, και επομένως $\alpha(G) \geq n - \beta(G)$. Το ζητούμενο προκύπτει συνδυάζοντας τις δύο ανισότητες. □

Μια άμεση συνέπεια της Πρότασης 2 είναι ότι ένα σύνολο ανεξαρτησίας S είναι μέγιστο αν και μόνο αν το σύνολο κάλυψης $V \setminus S$ είναι ελάχιστο.

Ο υπολογισμός ενός μέγιστου συνόλου ανεξαρτησίας σε γενικά γραφήματα είναι ένα πολύ δύσκολο πρόβλημα (από άποψη υπολογιστικής πολυπλοκότητας). Ένα μεγιστοτικό σύνολο ανεξαρτησίας προκύπτει εύκολα αν ξεκινήσουμε με ένα σύνολο ανεξαρτησίας S (π.χ. αρχικά το κενό σύνολο). Ενόσω υπάρχει κορυφή $v \in V \setminus S$ που δεν συνδέεται με καμία κορυφή του S, αντικαθιστούμε το S με το $S \cup \{v\}$ και συνεχίζουμε. Αυτή η επαναληπτική διαδικασία τερματίζει με ένα μεγιστοτικό σύνολο ανεξαρτησίας S (αν προσθέσω οποιαδήποτε κορυφή στο S, αυτό παύει να είναι σύνολο ανεξαρτησίας).

Κατ' αναλογία, ο υπολογισμός ενός ελάχιστου συνόλου κάλυψης είναι ένα πολύ δύσκολο πρόβλημα (από άποψη υπολογιστικής πολυπλοκότητας). Είναι όμως σχετικά απλό να υπολογίσουμε ένα σύνολο κάλυψης που έχει το πολύ $2\beta(G)$ κορυφές.

Υπολογίζουμε ένα μεγιστοτικό ταίριασμα M. Γνωρίζουμε ότι οι ελεύθερες κορυφές του M αποτελούν ένα σύνολο ανεξαρτησίας. Συνεπώς, οι κορυφές που είναι ταιριασμένες στο M αποτελούν ένα σύνολο κάλυψης. Έστω λοιπόν C το σύνολο κάλυψης που αποτελείται από τις ταιριασμένες κορυφές στο M. Είναι $|C| = 2|M|$ (για κάθε ακμή του M έχουμε τα δύο άκρα της στο C). Όμως είναι $\beta(G) \geq |M|$ γιατί κάθε σύνολο κάλυψης (του ελάχιστου συμπεριλαμβανομένου) περιλαμβάνει τουλάχιστον ένα από τα δύο άκρα κάθε ακμής του M. Διαφορετικά, θα η συγκεκριμένη ακμή του M θα ήταν ακάλυπτη. Συνεπώς, $|C| = 2|M| \leq 2\beta(G)$.

Είδαμε λοιπόν ότι για κάθε ταίριασμα M και κάθε σύνολο κάλυψης C, ισχύει ότι $|M| \leq |C|$. Ο λόγος είναι ότι το σύνολο κάλυψης πρέπει να περιέχει τουλάχιστον ένα από τα δύο άκρα κάθε ακμής του ταιριάσματος. Μάλιστα η ισότητα αποτελεί κριτήριο βελτιστότητας (optimality criterion) τόσο για ένα ταίριασμα όσο και για το αντίστοιχο σύνολο κάλυψης.

Πρόταση 3. *Έστω ένα ταίριασμα M και σύνολο κάλυψης C τέτοια ώστε $|M| = |C|$. Τότε το M αποτελεί ένα μέγιστο ταίριασμα και το C αποτελεί ένα ελάχιστο σύνολο κάλυψης.*

Απόδειξη. Έστω M^* ένα μέγιστο ταίριασμα και C^* ένα ελάχιστο σύνολο κάλυψης. Ισχύει ότι

$$|M| \leq |M^*| \leq |C^*| \leq |C|$$

Η πρώτη ανισότητα ισχύει γιατί το M^* είναι ένα μέγιστο ταίριασμα, η δεύτερη γιατί το μέγεθος κάθε συνόλου κάλυψης είναι μεγαλύτερο ή ίσο από το μέγεθος κάθε ταιριάσματος, και η τρίτη ανισότητα γιατί το C^* είναι ένα ελάχιστο σύνολο κάλυψης. Αφού υποθέσαμε ότι $|M| = |C|$, όλες οι παραπάνω ανισότητες πρέπει να είναι ισότητες. Έτσι $|M| = |M^*|$ και $|C^*| = |C|$. □

Η παραπάνω πρόταση λέει ότι όταν ένα ταίριασμα έχει το ίδιο μέγεθος με ένα σύνολο κάλυψης, τότε και τα δύο είναι βέλτιστα (δηλ. το ταίριασμα είναι μέγιστο και το σύνολο κάλυψης ελάχιστο). Όμως υπάρχουν πολλά γραφήματα που το ελάχιστο σύνολο κάλυψης είναι μεγαλύτερο από το μέγιστο ταίριασμα.

2 Σύνολα Κάλυψης σε Διμερή Γραφήματα

Στη συνέχεια θα αποδείξουμε ότι στα διμερή γραφήματα το μέγεθος του μέγιστου ταιριάσματος είναι πάντα ίσο με το μέγεθος του ελάχιστου συνόλου κάλυψης. Αυτό το αποτέλεσμα είναι γνωστό σαν Θεώρημα του König και αποτελεί ουσιαστικά μια εναλλακτική διατύπωση του Θεωρήματος του Hall (συχνά τα δύο Θεωρήματα αναφέρονται σαν Θεώρημα König-Hall).

Θεώρημα 1 (Θεώρημα του König). *Σε ένα διμερές γράφημα, ο αριθμός των ακμών στο μέγιστο ταίριασμα είναι ίσος με τον αριθμό των κορυφών στο ελάχιστο σύνολο κάλυψης.*

Απόδειξη. Έστω $G(X, Y, E)$ διμερές γράφημα, και έστω M ένα μέγιστο ταίριασμα στο G. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι το M δεν είναι X-τέλειο[1]. Έστω $W \subseteq X$ το σύνολο των ελεύθερων κορυφών του X.

Εφαρμόζουμε τη διαδικασία κατασκευής δέντρων εναλλακτικών μονοπατιών που περιγράφηκε στην απόδειξη του Θεωρήματος του Hall. Αυτή τη φορά ξεκινάμε από το σύνολο W των ελεύθερων κορυφών του X. Αρχικά $Y_0 = \emptyset$. Σε κάθε φάση i, $i = 0, 1, 2, \ldots$, θέτουμε $X_{i+1} = M(Y_i) \cup W$ και $Y_{i+1} = \Gamma(X_{i+1})$.

Αφού το M είναι μέγιστο ταίριασμα, δεν υπάρχει επαυξητικό μονοπάτι για το M (βλ. Θεώρημα του Berge). Συνεπώς, αυτή η διαδικασία δεν μπορεί να καταλήξει σε σύνολο Y_i που περιέχει ελεύθερη κορυφή του Y (βλ. επίσης απόδειξη του Θεωρήματος του Hall). Αφού δεν είναι δυνατόν να προστίθενται συνεχώς νέες κορυφές στο σύνολο Y_i, σε κάποια φάση καταλήγουμε σε ένα Y_i με όλες τις κορυφές του ταιριασμένες και $\Gamma(M(Y_i) \cup W) = Y_i$. Είναι $X_i = M(Y_i) \cup W$.

Θεωρώ το σύνολο κορυφών $C = Y_i \cup (X \setminus X_i)$. Παρατηρούμε ότι το σύνολο $X \setminus X_i$ περιέχει μόνο ταιριασμένες κορυφές (όλες οι ελεύθερες κορυφές ανήκουν στο W και έχουν συμπεριληφθεί στο X_i). Μάλιστα τα "ταίρια" των κορυφών του $X \setminus X_i$ είναι οι ταιριασμένες κορυφές του Y που δεν ανήκουν στο Y_i. Πράγματι, μια ταιριασμένη κορυφή του X ανήκει στο X_i αν και μόνο αν το ταίρι της ανήκει στο Y_i. Ισοδύναμα, μια ταιριασμένη κορυφή του X δεν ανήκει στο X_i αν και μόνο αν το ταίρι της δεν ανήκει στο Y_i. Συνεπώς, ο αριθμός των κορυφών του C είναι ίσος με τον αριθμό των ταιριασμένων κορυφών στο Y (ή ισοδύναμα στο X), δηλαδή ο αριθμός των κορυφών του C είναι ίσος με τον αριθμό των ακμών του M (τυπικά, $|C| = |M|$).

Χρειάζεται ακόμη να δείξουμε ότι το C είναι ένα σύνολο κάλυψης. Αφού $\Gamma(X_i) = Y_i$, δεν υπάρχει καμία ακμή μεταξύ των κορυφών του X_i και των κορυφών του $Y \setminus Y_i$. Με άλλα λόγια, το $(Y \setminus Y_i) \cup X_i$ είναι ένα σύνολο ανεξαρτησίας. Επομένως, το $C = Y_i \cup (X \setminus X_i)$ είναι ένα σύνολο κάλυψης.

Αφού το C είναι σύνολο κάλυψης και $|C| = |M|$, το C είναι ένα ελάχιστο σύνολο κάλυψης λόγω της Πρότασης 3. Άρα το μέγεθος του ελάχιστου συνόλου κάλυψης είναι ίσο με το μέγεθος του μέγιστου ταιριάσματος. $\square$

Το Θεώρημα του Hall υποδεικνύει έναν αποδοτικό αλγόριθμο για τον υπολογισμό ενός μέγιστου ταιριάσματος σε ένα διμερές γράφημα. Σε ένα διμερές γράφημα, ένα ελάχιστο σύνολο κάλυψης μπορεί να υπολογισθεί από ένα μέγιστο ταίριασμα με βάση το Θεώρημα του König. Επιπλέον, ένα μέγιστο σύνολο ανεξαρτησίας μπορεί να υπολογισθεί παίρνοντας τις κορυφές που δεν ανήκουν στο ελάχιστο σύνολο κάλυψης. Επομένως, τα προβλήματα του υπολογισμού ενός ελάχιστου συνόλου κάλυψης και ενός μέγιστου συνόλου ανεξαρτησίας λύνονται αποδοτικά σε διμερή γραφήματα (αν και αποτελούν δυσεπίλυτα προβλήματα για γενικά γραφήματα).

3 Αριθμοί Ramsey

Μπορεί να αποδειχθεί ότι για κάθε ζευγάρι ακεραίων n, m, υπάρχει ένας ελάχιστος αριθμός $r(n, m)$ τέτοιος ώστε κάθε γράφημα με τουλάχιστον $r(n, m)$ κορυφές περιέχει είτε το K_n (κλίκα με n κορυφές) είτε το $\overline{K}_m$ (σύνολο ανεξαρτησίας με m κορυφές). Οι αριθμοί αυτοί συμβολίζονται

[1] Αφού το γράφημα είναι διμερές, το X αποτελεί σύνολο κάλυψης γιατί το Y αποτελεί σύνολο ανεξαρτησίας. Αν το M ήταν X-τέλειο, θα είχαμε $|M| = |X|$ και το ζητούμενο έπεται ευθέως από την Πρόταση 3.

με $r(n, m)$ και ονομάζονται αριθμοί Ramsey. Ο αριθμός Ramsey $r(n, m)$ είναι ο ελάχιστος που εξασφαλίζει την παραπάνω ιδιότητα με την έννοια ότι υπάρχει τουλάχιστον ένα γράφημα με $r(n, m) - 1$ κορυφές που δεν περιέχει είτε το K_n είτε το $\overline{K}_m$. Ο ακριβής υπολογισμός των αριθμών Ramsey για μεγάλες τιμές των n, m αποτελεί ένα εξαιρετικά δύσκολο πρόβλημα για το οποίο δεν γνωρίζουμε μια γενική μέθοδο επίλυσης.

Στη συνέχεια θα αποδείξουμε ότι $r(3, 3) = 6$. Παρατηρούμε αρχικά ότι $r(3, 3) \geq 6$ επειδή ο κύκλος με 5 κορυφές έχει μέγιστη κλίκα και μέγιστο σύνολο ανεξαρτησίας μεγέθους 2. Για να δείξουμε την ισότητα, πρέπει να δείξουμε ότι κάθε γράφημα με 6 κορυφές περιέχει είτε κλίκα είτε σύνολο ανεξαρτησίας με 3 κορυφές. Η προσθήκη και άλλων κορυφών δεν μπορεί να επηρεάσει αυτή την ιδιότητα.

Πρόταση 4. *Κάθε γράφημα με 6 κορυφές περιέχει είτε το K_3 είτε το $\overline{K}_3$.*

Απόδειξη. Έστω ότι στο γράφημα υπάρχει κορυφή v με βαθμό μεγαλύτερο ή ίσο του 3, και έστω u_1, u_2, u_3 τρεις γείτονες της v. Αν δύο από τις u_1, u_2, u_3 συνδέονται με ακμή (π.χ. η u_1 με τη u_2), το τρίγωνο v, u_1, u_2 αποτελεί κλίκα με 3 κορυφές. Διαφορετικά, οι u_1, u_2, u_3 αποτελούν σύνολο ανεξαρτησίας με 3 κορυφές.

Αν όλες οι κορυφές του γραφήματος έχουν βαθμό μικρότερο ή ίσο του 2, θεωρούμε το συμπληρωματικό γράφημα. Αυτό περιέχει κορυφή με βαθμό μεγαλύτερο ή ίσο του 3, και επομένως περιέχει είτε το το K_3 είτε το $\overline{K}_3$. Αν το συμπληρωματικό γράφημα περιέχει το K_3 (αντίστοιχα, το $\overline{K}_3$), το αρχικό γράφημα περιέχει το $\overline{K}_3$ (αντίστοιχα, το K_3). □

4 Χρωματικός Αριθμός Γραφημάτων

Ένας έγκυρος *χρωματισμός* ενός γραφήματος ονομάζεται μια ανάθεση χρωμάτων στις κορυφές ώστε κάθε ζευγάρι κορυφών που συνδέεται με ακμή να έχει διαφορετικό χρώμα. Ο *χρωματικός αριθμός* ενός γραφήματος είναι ο ελάχιστος αριθμός χρωμάτων για τον οποίο υπάρχει ένας έγκυρος χρωματισμός. Ο χρωματικός αριθμός ενός γραφήματος G συμβολίζεται με $\chi(G)$.

Σε έναν έγκυρο χρωματισμό, οι κορυφές με το ίδιο χρώμα συγκροτούν ένα σύνολο ανεξαρτησίας αφού δεν υπάρχει καμία ακμή μεταξύ τους. Κάθε έγκυρος χρωματισμός διαμερίζει τις κορυφές του γραφήματος σε τόσα σύνολα ανεξαρτησίας όσα και τα χρώματα που χρησιμοποιεί. Ο χρωματικός αριθμός ενός γραφήματος είναι ο μικρότερος ακέραιος για τον οποίο οι κορυφές του μπορούν να διαμεριστούν σε σύνολα ανεξαρτησίας. Έτσι, ένα γράφημα έχει χρωματικό αριθμό μικρότερο ή ίσο του k αν και μόνο αν είναι k-μερές.

Τα διμερή γραφήματα έχουν χρωματικό αριθμό 2. Επομένως, ένα γράφημα έχει χρωματικό αριθμό 2 αν και μόνο αν δεν έχει κύκλους με περιττό μήκος. Σαν άμεση συνέπεια προκύπτει ότι ο κύκλος με n κορυφές, συμβολίζεται με C_n, έχει

$$\chi(C_n) = \begin{cases} 2 \text{ αν } n \text{ άρτιος} \\ 3 \text{ αν } n \text{ περιττός} \end{cases}$$

Επίσης είναι $\chi(K_n) = n$, $\chi(K_n - v) = n - 1$ για κάθε κορυφή v, και $\chi(\overline{K_n}) = 1$. Έχουμε ακόμη αποδείξει ότι κάθε επίπεδο γράφημα έχει χρωματικό αριθμό μικρότερο ή ίσο του 5. Μάλιστα το διάσημο Θεώρημα των 4 χρωμάτων λέει ότι κάθε επίπεδο γράφημα έχει χρωματικό αριθμό μικρότερο ή ίσο του 4.

Αποδεικνύεται εύκολα ότι κάθε γράφημα με μέγιστο βαθμό Δ έχει χρωματικό αριθμό το πολύ $\Delta + 1$ (βλ. Άσκηση 6, της 1ης Εργασίας). Η ιδέα είναι ότι τα $\Delta + 1$ χρώματα είναι αρκετά για να χρωματίσουμε μια κορυφή και τους γειτόνους της με διαφορετικά χρώματα. Επομένως, ο αλγόριθμος που εξετάζει τις κορυφές μία-προς-μία και χρωματίζει κάθε κορυφή με το μικρότερο διαθέσιμο χρώμα υπολογίζει ένα (έγκυρο) χρωματισμό των κορυφών με $\Delta + 1$ χρώματα το πολύ. Εναλλακτικά, αυτό μπορεί να αποδειχθεί εύκολα με μαθηματική επαγωγή.

Κάθε γράφημα που περιέχει μια κλίκα μεγέθους k έχει χρωματικό αριθμό τουλάχιστον k (χρειάζονται k διαφορετικά χρώματα για τις κορυφές της κλίκας). Γνωρίζουμε ότι η μεγαλύτερη κλίκα ενός γραφήματος G αντιστοιχεί στο μεγαλύτερο σύνολο ανεξαρτησίας του συμπληρωματικού γραφήματος $\overline{G}$. Επομένως, για κάθε γράφημα G, έχουμε $\chi(G) \geq \alpha(\overline{G})$.

Πρόταση 5. *Για κάθε γράφημα $G(V, E)$ με n κορυφές, $n/\alpha(G) \leq \chi(G) \leq n - \alpha(G) + 1$.*

Απόδειξη. Έστω $\chi(G) = k$. Οι κορυφές του G μπορούν να διαμεριστούν σε k σύνολα ανεξαρτησίας $V_1, V_2, \ldots, V_k$. Αφού το μεγαλύτερο σύνολο ανεξαρτησίας στο G έχει $\alpha(G)$ κορυφές, έχουμε $|V_i| \leq \alpha(G)$, για κάθε $i = 1, 2, \ldots, k$. Επομένως,

$$n = \sum_{i=1}^{k} |V_i| \leq k\,\alpha(G) = \chi(G)\alpha(G) \Rightarrow n/\alpha(G) \leq \chi(G)$$

Για το πάνω φράγμα, οι κορυφές του μεγαλύτερου συνόλου ανεξαρτησίας μπορούν να χρωματιστούν με ένα χρώμα. Ο χρωματισμός των υπόλοιπων $n - \alpha(G)$ κορυφών απαιτεί το πολύ $n - \alpha(G)$ διαφορετικά χρώματα. Συνολικά, ο χρωματικός αριθμός του G δεν μπορεί να ξεπερνά το $n - \alpha(G) + 1$. □

Άσκηση 1. Να δώσετε ένα παράδειγμα γραφήματος G με n κορυφές και χρωματικό αριθμό $\chi(G) = n/\alpha(G)$, και ένα παράδειγμα γραφήματος G' με n κορυφές και χρωματικό αριθμό $\chi(G') = n - \alpha(G') + 1$.