

אלגוריתמים לבעיות אופטימיזציה ברשתות ובגרפים

חיבור לשם קבלת תואר "דוקטור לפילוסופיה"
מאת

שי גוטנר

בהנחייתם של
פרופסור נוגה אלון ופרופסור יוסי עזר

הוגש לסנאט של אוניברסיטת תל-אביב
ינואר 2009

תמצית

ישנה ספרות ענפה העוסקת באלגוריתמים מקוונים ומקורבים לבעיות ברשתות. בעבודה זו אנו חוקרים מגוון בעיות אופטימיזציה, תוך התמקדות ב-admission control, מניח מקורבת של תת-גרפים קטנים בתוך גרפים גדולים, בעיית הקבוצה השלטת המינימאלית ובעיית יער השטיינר המוכללת עם קנסות.

תחילה אנו דנים בפרק 1 בבעיית admission control ברשתות כלליות. בקשות תקשורת מגיעות לאורך זמן, והאלגוריתם המקוון מקבל או דוחה כל בקשה תוך שמירה על מגבלות הקיבולת של הרשת. אנו בוחנים אלגוריתמים מקוונים שיכולים לדחות בקשות שכבר התקבלו, והמטרה שלהם היא להביא למינימום את מספר הבקשות שנידחו. כל בקשה מגיעה יחד עם המסלול עליו היא צריכה להיות מנותבת. אנו מראים אלגוריתם אקראי $O(\log^2(mc))$ -תחרותי למקרה הממושקל, כאשר m הוא מספר הקשתות בגרף ו- c היא הקיבולת המקסימאלית של קשת. למקרה הלא-ממושקל, אנו מציגים אלגוריתם אקראי $O(\log m \log c)$ -תחרותי. תוצאה זו נותנת מענה לבעיה פתוחה של Kalai, Blum ו-Kleinberg שהועלתה במאמר [37].

בעיית admission control הינה הכללה של הבעיה המקוונת של כיסוי בקבוצות עם חזרות, שהקלט שלה הוא משפחה של m תת-קבוצות של קבוצת בסיס בת n איברים. איברים מקבוצת הבסיס ניתנים לאלגוריתם המקוון בזה אחר זה, כאשר אפשרי שאיבר יידרש מספר פעמים. האלגוריתם מחויב לכסות כל איבר על-ידי תת-קבוצות שונות, בהתאם למספר הפעמים שהאיבר נדרש.

אנו נותנים אלגוריתם אקראי $O(\log m \log n)$ -תחרותי לבעיה המקוונת של כיסוי בקבוצות עם חזרות. תוצאה זו תואמת לחסם תחתון של $\Omega(\log m \log n)$ שניתן ע"י Feige ו-Korman [74] עבור יחס התחרותיות של כל אלגוריתם אקראי בעל זמן ריצה פולינומי, תחת ההנחה ש-BPP \neq NP.

בפרק 2 אנו מגדירים סוג חדש ומעניין של משפחות של פונקציות hash מושלמות. משפחה של פונקציות מ- $[n]$ ל- $[k]$ היא משפחה (ϵ, k) -מאוזנת של פונקציות hash, אם קיים קבוע T חיובי כך שלכל $K \subseteq [n]$ בגודל $|K|=k$, מספר הפונקציות במשפחה שהן חד-חד-ערכיות על K הוא בין $(1-\epsilon)T$ ל- $(1+\epsilon)T$. המשפחה היא k -מאוזנת באופן מושלם אם היא $(0, k)$ -מאוזנת.

אנו מראים שכל משפחה k -מאוזנת באופן מושלם היא בגודל לפחות $c(k)n^{\lfloor k/2 \rfloor}$, וכמו-כן לכל $\epsilon > 1/\text{poly}(k)$ קיימות בניות מפורשות של משפחות (ϵ, k) -מאוזנות של פונקציות hash מ- $[n]$ ל- $[k]$ בגודל $e^{(1+o(1))k} \log n$. תוצאה זו היא הדוקה עד כדי גורם ה- $o(1)$ שבחזקה, ומספקת אלגוריתם דטרמיניסטי פולינומי למניח מקורבת של מספר המסלולים או המעגלים באורך נתון k (או עותקים של כל גרף H עם k צמתים ו-tree-width חסום) בתוך גרף קלט בגודל n , עד כדי טעות יחסית ϵ , לכל $k \leq O(\log n)$.

מספר השליטה של גרף $G=(V,E)$ הוא הגודל המינימאלי של קבוצה שלטת $U \subseteq V$, המקיימת שכל צומת ב- $V-U$ הוא שכן של לפחות צומת אחד מ- U . בפרק 3 אנו נותנים אלגוריתם בעל זמן ריצה $n^{O(dk)}$ למציאת קבוצה שלטת בגודל לכל היותר k בגרף d -degenerated עם n צמתים. תוצאה זו מוכיחה שבעיית הקבוצה השלטת היא fixed-parameter tractable עבור גרפים שהם degenerated. לגרפים שלא מכילים את K_h כמינור טופולוגי, אנו נותנים אלגוריתם מהיר יותר עם זמן ריצה $n^{O(h)}$. לגרפים שהינם חסרי K_h -מינור, זמן הריצה משתפר עוד יותר עד כדי $n^{O(\log h)}$.

המושג של problem kernel מתייחס לאלגוריתם שרץ בזמן פולינומי ומפחית באופן מוכח את גודל הקלט של בעיה. בהינתן גרף G שמספר השליטה שלו הוא k , המטרה היא לתכנן אלגוריתם פולינומי שמייצר גרף G' שהגודל שלו תלוי רק ב- k , וגם לו יש מספר שליטה השווה ל- k .

אנו מציגים את התוצאה הלא-טריוויאלית הראשונה למקרה הכללי של גרפים עם מינור אסור, כדלהלן. לכל h קבוע, בהינתן גרף G עם n צמתים שלא מכיל את K_h כמינור טופולוגי, האלגוריתם שלנו מייצר בזמן $O(n^{3.5} + k^{O(1)})$ תת-גרף G' של G , כך שאם מספר השליטה של G הוא k , אז מספר השליטה של G' הוא גם כן k ול- G' יש לכל היותר k^c צמתים, כאשר c הינו קבוע התלוי רק ב- h . תוצאה זו משופרת לגרפים שלא מכילים $K_{3,h}$ כמינור טופולוגי, באמצעות אלגוריתם פשוט יותר שמייצר תת-גרף עם ck צמתים, כאשר c הוא קבוע התלוי רק ב- h . באמצעות הבנייה של kernel פולינומי, אנו נותנים אלגוריתם בעל זמן ריצה $O(n^{3.5} + 2^{O(nk)})$ למציאת קבוצה שלטת בגודל לכל היותר k בגרף חסר H -מינור עם n צמתים.

פרק 4 דן באלגוריתמי קירוב לבעיית יער השטיינר המוכללת עם קנסות, המוגדרת כדלהלן. הקלט הוא גרף לא מכוון $G=(V,E)$, אוסף $T=\{T_1, \dots, T_k\}$, שכל אחד ממנו הוא תת-קבוצה של V בגודל לפחות 2, פונקצית משקל $w:E \rightarrow \mathbb{R}^+$, ופונקצית קנס $p:T \rightarrow \mathbb{R}^+$. המטרה היא למצוא יער F המביא למינימום את העלות של הקשתות של F בתוספת הקנסות המשולמים עבור כל תת-קבוצה T_i שהצמתים שלה אינם מקושרים כולם על-ידי F . התוצאה המרכזית שלנו היא $(3-4/n)$ -קירוב לבעיית יער השטיינר המוכללת עם קנסות, כאשר $n \geq 2$ הוא מספר הצמתים בגרף. אלגוריתם הקירוב מתקבל באמצעות יישום של שיטת ה-local ratio, והוא פשוט הרבה יותר מהאלגוריתם הקומבינטורי הטוב ביותר הידוע לבעיה זו.

הגישה שלנו נותנת $(2-1/(n-1))$ -קירוב לבעיית עץ השטיינר עם קנסות (כל תת-הקבוצות T_i הן בגודל 2 ויש צומת שורש r ששייך לכולן). אלגוריתם קירוב זה הינו למעשה גרסת ה-local ratio לאלגוריתם ה-primal-dual של Goemans ו-Williamson מהמאמר [56]. מקרה פרטי נוסף של האלגוריתם המרכזי שלנו הוא אלגוריתם ה-local ratio של Bar-Yehuda שנותן $(2-2/n)$ -קירוב לבעיית יער השטיינר המוכללת (כל הקנסות הם אינסוף) [31].

תקציר

התמודדות עם NP-קשיות הינה המוטיבציה העיקרית לתוצאות האלגוריתמיות המוצגות בתזה זו. בעיות NP-קשות אינן ניתנות לפתרון על-ידי אלגוריתם בעל זמן ריצה פולינומי, אלא אם כן $P=NP$, והדבר מוביל באופן טבעי לחקירה של אלגוריתמים יעילים המחפשים פתרונות מקורבים. המטרה הן של אלגוריתמים מקוונים והן של אלגוריתמים מקורבים היא להשיג פתרונות שהינם קרובים ככל הניתן לפתרון האופטימאלי. הקושי העיקרי של אלגוריתם מקוון הינו העובדה שהוא חייב לקבל החלטות מבלי לדעת את כל הקלט, בעוד המגבלה העיקרית של אלגוריתם קירוב היא שהוא חייב לרוץ בזמן פולינומי.

גישה מעניינת נוספת להתמודדות עם NP-קשיות של בעיות ניתנת על-ידי התחום החדש של parameterized complexity. יעד המחקר העיקרי של תחום זה הוא תכנון אלגוריתמים מדויקים, שזמן הריצה שלהם פולינומי בגודל הקלט, אבל ייתכן שיהיה מעריכי בפרמטר מסוים של הבעיה. משמעות הדבר היא שהפרמטר, שהוא בדרך-כלל גודל הפתרון האופטימאלי, הוא הגורם המשפיע העיקרי על סיבוכיות הבעיה.

דרך אחרת לנתח את ה-parameterized complexity של בעיה היא באמצעות בניה של kernels. המושג של problem kernel מתייחס לאלגוריתם שרץ בזמן פולינומי ומפחית באופן מוכח את גודל הקלט של בעיה. ניתן להשתמש ב-problem kernels כדי לשפר את היעילות של אלגוריתמים מקורבים ומדויקים, והם שימושיים גם לבעיות מעשיות.

בתזה זו אנו מתכננים ומנתחים אלגוריתמים מקוונים ומקורבים למגוון בעיות אופטימיזציה ברשתות.

התזה מורכבת מארבעה פרקים. הפרק הראשון עוסק בבעיית ה-admission control ברשתות כלליות, כאשר המטרה היא להביא למינימום את מספר הדחיות שמתבצעות על-ידי האלגוריתם המקוון. בפרק השני אנו מציגים את המושג החדש של משפחות מאוזנות של פונקציות hash, ומדגימים את היישומים שלהן למנייה מקורבת של מספר העותקים של תת-גרף קטן בתוך גרף גדול. הפרק השלישי דן ב-parameterized complexity של בעיית הקבוצה השלטת בגרפים שהם degenerated ובגרפים עם מינור אסור. אנו חוקרים אלגוריתמים מדויקים שהם fixed-parameter tractable לבעיה זו, וכן את הבניה של problem kernels בגודל פולינומי. הפרק הרביעי והאחרון מוקדש למחקר של אלגוריתמי קירוב קומבינטוריים לבעיית יער השטיינר המוכללת עם קנסות.

ההמרה של אלגוריתמים אקראיים לאלגוריתמים דטרמיניסטיים התבררה כמשימה מורכבת במספר מקרים. תהליך הדרנדומיזציה היה כרוך בשימוש בטכניקות שונות, כולל שיטת התחלות המותנות, almost k-wise independence, קודים לתיקון שגיאות וגרפים מרחיבים.

האלגוריתמים שלנו לבעיית הקבוצה השלטת מתבססים על תוצאות קומבינטוריות לגבי גרפים שהם degenerated וגרפים עם מינור אסור. אחת המסקנות של תוצאות אלו היא אלגוריתם קירוב קבוע, שביחד עם טכניקות אחרות נותן אלגוריתם מדויק מהיר יותר לבעיית הקבוצה השלטת על גרפים חסרי H-מינור.

נתאר כעת את תוכנו של כל פרק ופרק.

פרק 1: בעיית ה-Admission Control תוך

מזעור דחיות

אנו חוקרים את בעיית ה-admission control בגרפים כלליים עם קיבולות לקשתות. אלגוריתם מקוון מקבל סידרה של בקשות תקשורת על מסלול וירטואלי, שניתן לקבל או לדחות, תוך שמירה על מגבלות הקיבולת.

אנו לומדים את המטרה של הבאה למינימום של מספר הבקשות שנידחו, שנחקרה לראשונה במאמר [37]. גישה זו מתאימה ליישומים בהם דחיות של בקשות אמורות להיות דבר נדיר. מצב שבו חלק משמעותי של הבקשות נידחה אפילו על-ידי הפתרון האופטימאלי משמעותו שצריך לשדרג את הרשת.

בפרק זה, אנו בוחנים אלגוריתמים מקוונים לבעיית ה-admission control, שיכולים לדחות בקשות שכבר התקבלו. כל בקשה מגיעה יחד עם המסלול עליו היא צריכה להיות מנותבת. האלגוריתם מחליט האם לקבל או לדחות את הבקשה.

אנו מגדירים עכשיו באופן פורמאלי את בעיית ה-admission control. הקלט מכיל את המרכיבים הבאים:

- גרף מכוון $G=(V,E)$, כאשר $|E|=m$. לכל קשת e יש קיבולת שלמה $c_e > 0$. אנו מסמנים $c = \max_{e \in E} c_e$.
- סידרה של בקשות r_1, r_2, \dots , שכל אחת מהן היא מסלול פשוט בגרף. לכל בקשה r_i יש עלות $p_i > 0$ המשויכת אליה.

פתרון אפשרי לבעיה חייב להבטיח שלכל קשת e , מספר הבקשות שהתקבלו שהמסלול שלהן מכיל את e הוא לכל היותר הקיבולת c_e . המטרה היא למצוא פתרון אפשרי כך שעלות הבקשות שנידחו הינה מינימאלית. האלגוריתם המקוון מקבל בקשה אחת כל פעם, וחייב להחליט האם לקבל או לדחות כל בקשה. מותר לדחות בקשה לאחר שכבר התקבלה, אך אסור לקבל בקשה שנידחתה קודם לכן.

יהא OPT הפתרון האפשרי בעל העלות המינימאלית C_{OPT} . אלגוריתם יקרא β -תחרותי אם העלות הכוללת של הבקשות שנידחו על-ידי האלגוריתם היא לכל היותר βC_{OPT} .

בעיית ה-admission control תוך מזעור דחיות נחקרה על-ידי Kalai, Blum ו-Kleinberg במאמר [37], שבו מוצגים שני אלגוריתמים דטרמיניסטיים עם יחסי תחרותיות של $O(\sqrt{m})$ ושל $c+1$ (הוא מספר הקשתות בגרף c ו- c היא הקיבולת המקסימאלית). הם מעלים את השאלה האם ניתן להשיג אלגוריתם מקוון עם יחס תחרותיות פולי-לוגריתמי.

בפרק זה אנו מראים שבעיית ה-admission control תוך מזעור דחיות היא הכללה של הבעיה המקוונת של כיסוי בקבוצות עם חזרות, המוגדרת כדלהלן: תהא X קבוצת בסיס עם n

איברים, ותהא S משפחה של תת-קבוצות של X , $|S|=m$. לכל $S \in \mathcal{S}$ יש עלות אי-שלילית המשויכת אליה. מנגנון יריב נותן לאלגוריתם איברים מתוך X מזה אחר זה. כל איבר של X יכול להינתן מספר כלשהו של פעמים, לא בהכרח ברצף. איבר צריך להיות מכוסה על-ידי קבוצות שמספרן שווה למספר הפעמים שהוא הופיע בסדרת הבקשות (לכל קבוצה יש רק עותק אחד). אנו מניחים שהקבוצות X ו- S ידועות לאלגוריתם כבר בהתחלה, אולם האיברים הניתנים על-ידי המנגנון היריב אינם ידועים מראש. המטרה היא למזער את העלות של הקבוצות שנבחרות על-ידי האלגוריתם.

התוצאה העיקרית שאנו מציגים בפרק זה היא אלגוריתם אקראי $O(\log^2(mc))$ -תחרותי לבעיית ה-admission control תוך מזעור דחיות. הדבר פותר את השאלה הפתוחה שהועלתה על-ידי Blum et al. [37]. למקרה הלא-ממושקל, כשכל העלויות שוות ל-1, אנו משרים במקצת חסם זה ונותנים אלגוריתם אקראי $O(\log m \log c)$ -תחרותי.

אנו מציגים רדוקציה פשוטה בין בעיית הכיסוי בקבוצות עם חזרות ובעיית ה-admission control. מכאן נובע אלגוריתם אקראי $O(\log^2(mn))$ -תחרותי לבעיה המקוונת של כיסוי בקבוצות עם חזרות. למקרה הלא-ממושקל (כל העלויות שוות ל-1), אנו נותנים אלגוריתם אקראי $O(\log m \log n)$ -תחרותי. תוצאה זו תואמת לחסם התחתון של $\Omega(\log m \log n)$ שניתן על-ידי Feige ו-Korman [74]. מהתוצאות שלהם נובע גם חסם תחתון של $\Omega(\log m \log c)$ ליחס התחרותיות של כל אלגוריתם אקראי פולינומי לבעיית ה-admission control תוך מזעור דחיות (תחת ההנחה $BPP \neq NP$).

מקורות: התוצאות של פרק זה מופיעות במאמר:

- N. Alon, Y. Azar and S. Gutner, Admission control to minimize rejections and online set cover with repetitions, Proc. of the 17th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA 2005), 238-244.
Also: ACM Transactions on Algorithms, to appear.

פרק 2: משפחות מאוזנות של פונקציות Hash

הטכניקה האלגוריתמית Color Coding מאפשרת להכריע באופן יעיל האם גרף קלט נתון מכיל מסלול או מעגל באורך מסוים, או כל תת-גרף מוגדר בעל tree-width חסום. כאשר מתמקדים, לשם פשטות, במסלולים, השיטה נותנת אלגוריתם דטרמיניסטי כדי להכריע, בזמן $O(k \log |V|)$, האם גרף (מכוון או לא מכוון) $G=(V,E)$ מכיל מסלול (פשוט) על k צמתים. הגישה הבסיסית, שהוצגה לראשונה במאמר [22], הינה פשוטה מאוד. תחילה היא נותנת אלגוריתם אקראי, ואז ממירה אותו לאלגוריתם דטרמיניסטי. האלגוריתם האקראי צובע תחילה את הצמתים של G באמצעות k צבעים באופן אקראי. נקרא למסלול על k צמתים (k -מסלול, בקיצור) צבעוני אם הצמתים שלו נצבעו בכל k הצבעים השונים. לא קשה לבדוק בזמן $O(k^2|E|)$, באמצעות תכנות דינאמי, האם יש מסלול צבעוני. מאחר שההסתברות של k -מסלול להפוך להיות צבעוני בצביעה מקרית היא $e^{-k} < k!/k^k$, חזרה על תהליך זה Ce^k פעמים מספקת אלגוריתם אקראי שבו ההסתברות לא למצוא מסלול שאכן קיים הינה קטנה מ- e^{-C} . העניין המכריע בדרנדומיזציה של אלגוריתם זה הינה ההבחנה שבניות ידועות של משפחות של פונקציות hash הניתנות במאמר [79] בעקבות מאמר [55], נותנות משפחה מפורשת של $2^{O(k)} \log |V|$ צביעות של הצמתים של G ב- k צבעים, כך שהאיברים של כל קבוצה של k צמתים מקבלים צבעים שונים בלפחות אחת מהצביעות. לכן, ניתן פשוט להריץ את אלגוריתם התכנות הדינאמי עבור כל אחת מהצביעות, ולקבל אלגוריתם דטרמיניסטי לבעיה.

לטכניקה לעיל נמצאו לאחרונה יישומים בביולוגיה חישובית (ראה [80], [81], [83], [64]). יישומים אלה מובילים לדיון בבעיה של מנייה, או קירוב של מספר ה- k -מסלולים (או גרפים אחרים עם tree-width חסום) בגרף נתון. מאחר שבאמצעות תכנות דינאמי קל לספור בדיוק את מספר ה- k -מסלולים הצבעוניים בגרף נתון, הקיום של אלגוריתם קירוב אקראי יעיל נובע די בקלות משימוש באותה גישה; הדבר נעשה במאמר [13].

כדי לבצע דרנדומיזציה של הליך המנייה האקראית (או המנייה המקורבת), יש צורך להכליל את המושג הרגיל של פונקציות hash. אנו עושים זאת באמצעות ההגדרה הבאה.

משפחה של פונקציות מ- $[n]$ ל- $[q]$ היא משפחה (ϵ, k) -מאוזנת של פונקציות hash, אם לכל $|S|=k, S \subseteq [n]$, מספר הפונקציות שהן חד-חד-ערכיות על S הוא בין $(1-\epsilon)T$ ל- $(1+\epsilon)T$ עבור קבוע

$T > 0$. המשפחה היא k -מאוזנת באופן מושלם אם היא $(0, k)$ -מאוזנת, היינו היא (ε, k) -מאוזנת עבור $\varepsilon = 0$.

במאמר [52] מוכיחים Flum ו-Grohe שהבעיה של מניה מדויקת של מספר המסלולים והמעגלים באורך k הן בגרפים מכוונים והן בגרפים לא מכוונים, כבעיה שהפרמטר שלה הוא k , הינה שלמה ב- $\#W[1]$. מכאן נובע שלא סביר שיש אלגוריתם בעל זמן ריצה $f(k) \cdot n^c$ למניה מדויקת של מספר המסלולים או המעגלים באורך k בגרף בגודל n לכל פונקציה ניתנת לחישוב $f: N \rightarrow N$ וקבוע c .

למרות זאת, הבעיה של קירוב מספרים אלה הינה נגישה יותר. במאמר [24] השיגו Arvind ו-Raman אלגוריתם אקראי *fixed-parameter tractable* למניה מקורבת של מספר העותקים של k -מסלולים (או כל תת-גרף קבוע עם *tree-width* חסום) בתוך גרף גדול. אלגוריתם קירוב דומה מופיע במאמר [13].

התוצאות של Flum ו-Grohe שהוזכרו לעיל מצביעות על כך לא קיימת משפחה k -מאוזנת באופן מושלם של פונקציות *hash* מ- $[n]$ ל- $[k]$ שגודלה $n^{O(1)}$. $f(k)$. אנו מראים תוצאה חזקה יותר, על-ידי הוכחה שכל משפחה k -מאוזנת באופן מושלם של פונקציות *hash* מ- $[n]$ ל- $[q]$ היא בגודל לפחות $n^{\lfloor k/2 \rfloor}$, כאשר $c(k, q)$, $c(k, q)$ הוא קבוע חיובי התלוי רק ב- k ו- q . אנו מבחינים בכך שתוצאה זו לא רחוקה מלהיות הדוקה, שכן לכל $n > k$ קיימת משפחה k -מאוזנת באופן מושלם של פונקציות מ- $[n]$ ל- $[k]$ בגודל לכל היותר n^{k-1} . הדבר מראה שגישה ה-Color Coding לא יכולה לספק אלגוריתם לספירת k -מסלולים בגרף עם n צמתים בזמן $o(n^{\lfloor k/2 \rfloor})$.

התוצאה החיובית המרכזית שלנו היא בניה מפורשת, לכל $1/\text{poly}(k) < \varepsilon \leq 1$, של משפחה (ε, k) -מאוזנת של פונקציות *hash* מ- $[n]$ ל- $[k]$ בגודל $n \log n \cdot e^{(1+o(1))k}$. זמן הריצה של הפרוצדורה שמספקת את הבניה הוא $n \log n \cdot e^{(1+o(1))k}$. יש לשים לב לכך שגודל המשפחה הינו אופטימאלי עד כדי גורם השגיאה $o(1)$, במעריך, שכן יש חסם תחתון ידוע של $\Omega(e^k \log n / \sqrt{k})$ לגודל של כל משפחה של פונקציות *hash* מ- $[n]$ ל- $[k]$ (גם אם היא לא מאוזנת והדרישה היחידה היא שכל קבוצה בגודל k ממופה באופן חד-חד-ערכי לפחות פעם אחת).

קיבלנו אם כן אלגוריתמי קירוב דטרמיניסטיים למניה של מספר ה- k -מסלולים הפשוטים בגרף $G=(V, E)$ עד כדי שגיאה יחסית של $\varepsilon = 1/\text{poly}(k)$ בזמן $|E| \log |V| \cdot 2^{O(k)}$. תוצאות דומות תקפות למניה מקורבת של מספר העותקים של כל גרף בגודל k עם *tree-width* קבוע. יש לשים לב לכך שביטוי זה הינו פולינומי לכל $k \leq O(\log n)$, ולא סביר שניתן להשיג תוצאה טובה יותר, שכן

דבר זה יגרור קיום אלגוריתם בעל זמן ריצה $2^{o(n)}$ לבעיית המסלול ההמילטוני, בסתירה להשערת הזמן המעריכי של [65,66].

מקורות: התוצאות של פרק זה מופיעות במאמרים:

- N. Alon and S. Gutner,
Balanced families of perfect hash functions and their applications,
Proc. of the 34th International Colloquium on Automata, Languages and
Programming (ICALP 2007), 435-446.
Also: ACM Transactions on Algorithms, to appear.
- N. Alon and S. Gutner,
Balanced hashing, color coding and approximate counting, submitted.

פרק 3: בעיית הקבוצה השלטת על גרפים עם

מינור אסור

הקלט לבעיה פרמטרית הוא זוג (x,k) , כאשר x הוא מופע הבעיה, k הוא הפרמטר, וגודל הקלט הוא $n:=|(x,k)|$. בעיה פרמטרית היא fixed-parameter tractable אם ניתן לפתור אותה בזמן $f(k) \cdot n^c$, עבור פונקציה ניתנת לחישוב $f:N \rightarrow N$ וקבוע c . אלגוריתם למציאת kernel הוא פונקציה הניתנת לחישוב בזמן פולינומי שבהינתן קלט (x,k) מחזירה קלט שקול (x',k') , כך שמתקיים $k' \leq k$ וגם $|x'| \leq g(k)$ עבור פונקציה ניתנת לחישוב $g:N \rightarrow N$. התמונה x' נקראת ה-problem kernel של x . בפרק זה, המושג של problem kernel עבור בעיית הקבוצה השלטת מתייחס לאלגוריתם פולינומי שבהינתן גרף G שמספר השליטה שלו הוא k , בונה גרף G' שהגודל שלו תלוי רק ב- k , וגם לו מספר שליטה השווה ל- k .

ידוע וקל להוכיח שלבעיה פרמטרית יש problem kernel אם ורק אם היא fixed-parameter tractable. לכן, אלגוריתם fixed-parameter לבעיית הקבוצה השלטת נותן kernel טריוויאלי שגודלו הוא פונקציה מסוימת של k , לאו דווקא פולינום. ניתן להשתמש ב-problem kernels כדי לשפר את היעילות של אלגוריתמים מקורבים ומדויקים, והם שימושיים גם לבעיות מעשיות.

ידוע שבעיית הקבוצה השלטת על גרפים כלליים הינה שלמה ב-[2] W [47]. מכאן נובע שלא סביר שיש אלגוריתם בעל זמן ריצה $f(k) \cdot n^c$ למציאת קבוצה שלטת בגודל לכל היותר k בגרף עם n

צמתים לכל פונקציה ניתנת לחישוב $f: N \rightarrow N$ וקבוע c . הדבר מוביל למחקר של משפחות של גרפים עבורם הבעיה היא fixed-parameter tractable.

התוצאה העיקרית של הפרק היא שבעיית הקבוצה השלטת היא fixed-parameter tractable עבור גרפים שהם degenerated. זמן הריצה המתקבל הוא $n^{O(dk)}$ למציאת קבוצה שלטת בגודל k בגרף d -degenerated עם n צמתים. האלגוריתם הינו ליניארי במספר הצמתים של הגרף, ואנו משפרים את התלות ב- k עובר המקרים הפרטיים הבאים. עבור גרפים שאינם מכילים את K_h כמינור טופולוגי, ניתן אלגוריתם משופר עם זמן ריצה של $(O(h))^{hk}$. עבור גרפים חסרי K_h -מינור, זמן הריצה שמושג הוא $(O(\log h))^{hk/2}$. אנו מראים שניתן להכליל את כל האלגוריתמים למקרה הממושקל במובן הבא. ניתן באותו זמן ריצה למצוא קבוצה שלטת בעלת משקל מינימאלי שגודלה לכל היותר k .

כללי הרדוקציה שהוצגו על-ידי Alber, Fellows ו-Niedermeier היו הראשונים להשיג problem kernel ליניארי עבור גרפים מישוריים [6]. גודל ה-kernel שהושג היה $335k$, כאשר k הינו מספר השליטה של הגרף. במאמר [54] הוכיחו Fomin ו-Thilikos שכללי הרדוקציה של Alber et al. נותנים kernel ליניארי בגודל $O(k+g)$ עבור גרפים עם גנוס g . במאמר [42] שיפרו Chen et al. את החסם העליון עבור המקרה המישורי ל- $67k$. הם הראו גם את החסם התחתון הראשון, בכך שהוכיחו שלכל $\epsilon > 0$, אין kernel בגודל $(2-\epsilon)k$ לבעיית הקבוצה השלטת על גרפים מישוריים, אלא אם כן $P=NP$.

באמצעות הצגת כלל רדוקציה מקורי, אנו מוכיחים שלבעיית הקבוצה השלטת על גרפים עם מינור אסור יש problem kernel בגודל פולינומי. תוצאה זו נותנת אלגוריתם בעל זמן ריצה $O(n^{3.5} + 2^{O(\sqrt{k})})$ למציאת קבוצה שלטת בגודל לכל היותר k בגרף חסר H -מינור עם n צמתים. לגרפים חסרי $K_{3,h}$ -מינור, אנו מראים שכללי הרדוקציה של Alber, Fellows ו-Niedermeier [6] נותנים kernel בגודל ליניארי. לכל כללי הרדוקציה המתוארים בפרק יש את התכונה שהשינויים היחידים שהם עושים לגרף קלט הם מחיקה של צמתים וקשתות. פירוש הדבר שהגרף המתקבל, כתוצאה מהפעלת החוקים, הוא תת-גרף של גרף הקלט. היתרונות של גישה זו הם הפשטות שלה והשמירה על תכונות מונוטוניות, כמו מישוריות, חוסר ב- H -מינור ו-degeneracy. אנו מראים שניתן להציג את הכללים של Alber et al. בצורה זו.

אנו דנים בשתי שאלות פתוחות שהוצגו על-ידי Cai, Chan ו-Chan במאמר [41] באשר לאלגוריתמים ליניאריים למציאת מעגל מושרה בגרפים שהם degenerated. מוצג אלגוריתם

אקראי בעל תוחלת זמן ריצה $O(n)$ למציאת מעגל מושרה באורך k בגרף עם מינור אסור על n צמתים. אנו משפרים את הדרנדומיזציה שמתבצעת במאמר [41] כדי לקבל אלגוריתם דטרמיניסטי בעל זמן ריצה $O(n \log n)$ לבעיה זו. באשר למציאת מעגלים מושרים בגרפים שהם degenerated, אנו מראים אלגוריתם דטרמיניסטי בעל זמן ריצה $O(n)$ למציאת מעגלים באורך לכל היותר 5, ומסבירים מדוע בלתי סביר שניתן להשיג תוצאה זו עבור מעגלים ארוכים יותר.

מקורות: התוצאות של פרק זה מופיעות במאמרים:

- N. Alon and S. Gutner,
Linear time algorithms for finding a dominating set of fixed size in degenerated graphs,
Proc. of the 13th Annual International Computing and Combinatorics Conference (COCOON 2007), 394-405.
Also: Algorithmica, to appear.
- N. Alon and S. Gutner,
Kernels for the dominating set problem on graphs with an excluded minor,
Electronic Colloquium on Computational Complexity (ECCC),
Report TR08-066, 2008.

פרק 4: בעיות עץ שטיינר עם קנסות

ישנה ספרות ענפה העוסקת באלגוריתמי קירוב לבעיות עץ שטיינר עם קנסות. מטרת פרק זה היא להציג אלגוריתמים קומבינטוריים אלגנטיים לבעיות אלו. טכניקת ה-local ratio [30,31,32] שאנו משתמשים בה מאפשרת לנו לתכנן אלגוריתמים פשוטים שהניתוח שלהם הינו ישיר וקל.

פרק זה מתמקד בבעיית יער השטיינר המוכללת עם קנסות, שתסומן בקיצור PCGSF ומוגדרת כדלהלן. הקלט הוא גרף לא מכוון $G=(V,E)$, אוסף $T=\{T_1, \dots, T_k\}$, שכל אחד ממנו הוא תת-קבוצה של V בגודל לפחות 2, פונקציית משקל $w:E \rightarrow \mathbb{R}^+$, ופונקציית קנס $p:T \rightarrow \mathbb{R}^+$. המטרה היא למצוא יער F המביא למינימום את העלות של הקשתות של F בתוספת הקנסות המשולמים עבור כל תת-קבוצה T_i שהצמתים שלה אינם מקושרים כולם על-ידי F . מכאן שכל הצמתים של תת-הקבוצה T_i חייבים להיות באותו רכיב קשירות של F כדי להימנע מהקנס.

המקרה הפרטי של בעיית ה-PCGSF, שנקרא בעיית יער השטיינר עם קנסות (כל תת-הקבוצות T_i הן בגודל 2) זכה לאחרונה להתייחסות רבה. התאמה של אלגוריתם עיגול LP מהמאמר [35] נותנת 3-קירוב לבעיה זו. תוצאה זו שופרה במאמר [60] כדי לקבל אלגוריתם מבוסס LP שהוא 2.54-קירוב, ובנוסף לכך מוצג גם אלגוריתם קומבינטורי primal-dual שהוא 3-קירוב ומשתמש בלמה של Farkas.

מסגרת לבעיות מהסוג הנדון עם דרישות קישוריות כלליות ופונקציה תת-מודולרית להגדרת הקנסות נחקרה במאמר [82]. המודל שלהם ישים לבעיית ה-PCGSF המוגדרת בפרק זה וכן לבעיות המוצגות במאמרים [62,60]. המחברים נותנים אלגוריתם מסובך המשיג 3-קירוב תוך שימוש בטכניקת ה-primal-dual, יחד עם אלגוריתם עיגול LP שמשיג יחס קירוב של 2.54.

התוצאה המרכזית שלנו היא אלגוריתם local-ratio שנותן $(3-4/n)$ -קירוב לבעיית יער השטיינר המוכללת עם קנסות, כאשר $n \geq 2$ הוא מספר הצמתים בגרף. מכאן נובע כמובן אותו יחס קירוב למקרה הפרטי של בעיית יער השטיינר עם קנסות, שנחקרה במאמרים [60,58]. אנו מציגים גם אלגוריתם $(2-1/(n-1))$ -קירוב לבעיית עץ השטיינר עם קנסות (כל תת-הקבוצות T_i הן בגודל 2 ויש צומת שורש r ששייך לכולן). אלגוריתם זה הינו למעשה גירסת ה-local ratio לאלגוריתם ה-primal-dual של Goemans ו-Williamson מהמאמר [56]. מקרה פרטי נוסף של האלגוריתם המרכזי שלנו הוא אלגוריתם ה-local ratio של Bar-Yehuda שנותן $(2-2/n)$ -קירוב לבעיית יער השטיינר המוכללת (כל הקנסות הם אינסוף) [31]. לפיכך, תרומה חשובה של פרק זה היא בהצגת הכללה טבעית של המסגרת שהוצגה על-ידי Goemans ו-Williamson, ולאחר מכן על-ידי Bar-Yehuda.

מקורות: התוצאות של פרק זה מופיעות במאמר:

- S. Gutner,
Elementary approximation algorithms for prize collecting Steiner tree problems,
Proc. of the 2nd Annual International Conference on Combinatorial
Optimization and Applications (COCOA 2008), 246-254.
Also: Information Processing Letters 107 (2008), 39-44.