אלגוריתמים מבוזרים - הרצאה 7, פירוק רשתות

שרון מלטר, אתגר 17

22 בספטמבר 2024

תוכן עניינים

1 הקדמה

ניזכר בבעיית MIS. נסתכל על מקרה מיוחד שלה; קוטר גרף גדול מאוד.
יעילות הגרף לא נהדרת במודל שלמדנו...

במקום פשוט לפעול על כל הגרף, נחלק את הגרף למקבצים (קלאסטרים) ונפעיל עליהם את האלגוריתם. אם נוכל לעשות זאת ככה שלכל היותר $1/2$ מהצמתים לא נמצאים בקלאסטר בכל פעם, נוכל לאחר $logn$ איטרציות לחלק את כולם לקלאסטרים ובכך לייעל את האלגוריתם.

2 הגדרות

- $Network\ Decomposition$ (פירוק רשתות) -
חלוקה של צמתים ל- C קלאסטרים בעלי קוטר שאורכו לכל היותר D, כך שכל קלאסטר צבוע בצבע שונה מבין C צבעים, וקלאסטרים שכנים לא צבועים באותו צבע.
- פירוק רשתות עם קוטר חלש -
אלו קלאסטרים עם קוטר חלש שאורכו לכל היותר D. כלומר, המרחק בין כל שני צמתים בקלאסטר הוא לכל היותר D **בגרף הבסיס**. כלומר, המרחק בקלאסטר עצמו **יכול** להיות גדול יותר.

3 אפליקציה - MIS עם פירוק רשתות

בהינתן פירוק רשת עם קוטר חלש לגרף G עם C צבעים וקוטר D לקלאסטרים.
נראה כיצד ניתן לפתור MIS עבור G בזמן $O(CD)$.

האלגוריתם:

1. באיטרציה ה- i, בכל הקלאסטרים שצבועים בצבע i, נפתור MIS באלגוריתם שלמדנו לכל הצמתים בהם שאין להם שכן שכבר נמצא ב- MIS.

נשים לב שניתן להריץ את האלגוריתם במקביל בכל הקלאסטרים באותו הצבע, מכיוון שיש ביניהם **סימטריה**. אין ביניהם שכנים כך שהם אינם משפיעים זה על זה.

<u>סיבוכיות זמן:</u> יש לנו C איטרציות, אחת לכל צבע של קלאסטרים, כך שכל אחת דורשת $O(D)$ זמן לאסיפת הקלאסטרים.
בניתוח יותר עדין נקבל שאנו צריכים $O(T+CD)$ זמן, כאשר T הוא זמן בניית פירוק הגרף. באופן דומה ניתן לפתור $(\Delta+1)$-צביעה ו- $(1+\epsilon)$-קירוב לכיסוי בצמתים ו- MDS.

4 בניית פירוק - (כי את 'מבוא ללוגיקה' עוד לא סיימנו)

כי איבדנו כל קשר למציאות.

להלן אלגוריתמים לבניית פירוק רשת:

1. בנייה רנדומית עבור $C,\ D=O(logn),\ T=O(log^2n)$.
2. בנייה דטרמיניסטית קלאסית עבור $C,\ D,\ T=2^{O(\sqrt{logn})}$.
האלגוריתם זה היה האלגוריתם הדטרמיניסטי המהיר ביותר ל- MIS, צביעה, MDS וכו'
3. אלגוריתם דטרמיניסטי שבונה פירוק $C,\ D,\ T=poly(logn)$.
זהו אלגוריתם דטרמיניסטי אשר מספק פיתרון בזמן פולינומי ללוגריתם עבור MIS, צביעה, MDS וכו'

נעבור לדבר על שיטות בנייה באופן מורחב יותר. נתחיל מגישה סדרתית לבנייה ונעבור לגישה מבוזרת.

5 בנייה סדרתית - *Sequential Construction*

לאלגוריתם החדש שנכיר יש $O(logn)$ איטרציות ובכל אחת נפלוט קבוצת קלאסטרים זרים בעלי קוטר $O(logn)$, כך שהקבוצה מגדירה מחלקת צבע.
ישנן $O(logn)$ איטרציות, כלומר $O(logn)$ צבעים בסך הכל.

5.1 האלגוריתם

נתאר את תהליך יצירת קבוצת קלאסטרים של צבע מסוים;

1. איטרציה ראשונה:
מתחילים בבחירת צומת u. לאחר מכן מגדילים סביבו 'כדור'. מגדילים את הכדור באופן הבא; בכל פעם r הוא המרחק המקסימלי של צומת בכדור מ־ r. נגדיל את r עד אשר מתקיים $|B(u,\ r+1)| \leqslant 2 \cdot |B(u,r)|$, כאשר $B(u,\ r)$ הוא גודל הכדור (מספר הצמתים בו).
נוסיף את $B(u,\ r)$ לסט הקלאסטרים ונמחק את $B(u,\ r+1)$ מהגרף.
ניתן להסיק כי $r = O(logn)$, מכיוון שעד ה־ r שנכפל, הכדור הכפיל את גודלו. הוא יכול לעשות זאת רק עד $O(logn)$ פעמים, מכיוון שיש בסך הכל n צמתים.

2. לאחר מכן, נבחר צומת חדש u' ונבצע את אותו התהליך עם הגרף שנותר.

3. נמשיך עם שלב 2 עד אשר הגרף ריק.

לבסוף ממשיכים את האלגוריתם על הצמתים שנמחקו אך לא שייכים לקלאסטר.

נוכיח שכל הקלאסטרים שקיבלנו בכל איטרציה הם זרים ולא שכנים, ובעלי קוטר $O(logn)$.

5.2 הוכחת נכונות

1. לאחר הוספת קלאסטר $B(u,\ r)$, מוחקים את $B(u,\ r+1)$ מהגרף. לכן הקלאסטר הבא מאותו הצבע לא יכול להכיל צמתים שהם שכנים של צמתי קלאסטר דומה מאותו הצבע. כמו כן הוא לא יכול להכיל צמתים של קלאסטר קודם.

2. הראנו כי $r = O(long)$ לפי בחירת r באלגוריתם ולכן הקוטר של כל קלאסטרים הוא $O(logn)$.

5.3 יעילות

למה: לאחר איטרציה אחת (יצירת קבוצה של צבע אחד), חצי מהצמתים בגרף שייכים לקלאסטר.

למה זו גוררת שנצטרך בסך הכל $O(logn)$ איטרציות לבנייה.

הוכחה:
לפי האלגוריתם, שבו בכל פעם מוסיפים קלאסטר לצבע עד אשר הגרף ריק, הצמתים היחידים בסוף האיטרציה שאינם בקלאסטר הינם הצמתים שב־ $B(u,\ r+1)$ ולא ב־ $B(u,\ r)$.
בחרנו r כך ש־ $|B(u,\ r+1)| \leqslant 2 \cdot |B(u,\ r)|$, ולכן בכל פעם שמוסיפים קלאסטר לקבוצה, לפחות חצי מהצמתים שנמחקו נמצאים בקלאסטר. לכן לפחות חצי מהצמתים בגרף נמצאים בקלאסטר בסוף האיטרציה.

עד כאן עם האלגוריתם הסדרתי שלנו.
אבל לא על זה שילמנו כסף- מה לגבי אלגוריתם מבוזר לפתרון הבעיה?

6 בנייה מבוזרת

משפט: קיים אלגוריתם מבוזר דטרמיניסטי שבונה פירוק רשת עם $C = O(logn)$ וקוטר חלש $D = O(log^3 n)$, הפועל ב־ $poly(logn)$ סיבובים.

עד כה אלגוריתם הבנייה היעיל ביותר שדיברנו עליו רץ ב־ $2^{\sqrt{logn}}$ סיבובים, כך שמדובר בשיפור משמעותי (!)

אם משפט זה נכון, אז נוכל להשתמש באלגוריתם שרץ ב־ $poly(logn)$ סיבובים לבעיות מרכזיות רבות, כגון MIS ו־ $(\Delta+1)$־צביעה.

6.1 האלגוריתם

נרצה שוב אלגוריתם שבכל איטרציה מעניק לנו קבוצת קלאסטרים לצבע, כך שבסוף האיטרציה לפחות חצי מהצמתים נמצאים בקלאסטר. כמו כן נרצה לכל היותר $O(logn)$ איטרציות.
אבל עד כאן כל אלו הן רק רשימת משאלות. איך מתחילים לבנות אלגוריתם כזה?

רעיון: נרצה $O(logn)$ שלבים, לכל צבע, כך שבשלב i, אין אף קשתות בין קלאסטרים ששונים בביט ה־ i (מימין) של המזהה שלהם.
נממש את המזהים כך:

1. לכל צומת ישנו מזהה, שזהו מזהה הקלאסטר שלהם. אזי ישנם $b = O(logn)$ במזהים (ולפי הרעיון גם b שלבים באלגוריתם)
2. בהתחלה, זהו המזהה של הצומת עצמו, כך שכל אחד "שייך לקלאסטר משל עצמו".

נתאר את האלגוריתם:

1. אם יש קשת בין שני צמתים בעלי ביט i שונה, הם יינסו לצרף את הקלאסטר של אחד לקלאסטר של השני.
2. השכן מוכן לקבל צומת נוסף לקלאסטר שלו רק אם הוא יגדל פי $(1 + \frac{1}{O(logn)})$ בעבור הוספה זו (בשלבים הבאים ההוספה כוללת את שאר הצמתים של אותו הקלאסטר שמבקש להתווסף)
3. אם קלאסטר לא מסכים להוספה, כל שכניו מקלאסטרים שונים נמחקים כך שהוא לא יכול לגדול יותר.

6.2 הוכחת נכונות

טענה: בכל שלב, כל הקלאסטרים מפסיקים לגדול ב־ $O(log^2 n)$ צעדים.

הוכחה:
לפי הגדרת האלגוריתם, קלאסטרים גדלים רק אם הם יכולים לגדול פי $1+\frac{1}{O(logn)}$. מאחר ש־ $(1+\frac{1}{O(logn)})^{O(log^2 n)} > n$, הקלאסטרים מפסיקים לגדול לאחר $O(log^2 n)$ צעדים.

בסך הכל ישנם $b = O(logn)$ שלבים, לכן, כפי שרצינו, הקוטר של כל קלאסטר הוא $poly(logn)$:)

טענה: לאחר i שלבים, ישנן קשתות רק בין קלאסטרים שמסכימים על i הביטים האחרונים שלהם.

הוכחה: באינדוקציה :)
בסיס: ניתן להראות שאין אף קשת בין צמתים בעלי ביט אחרון שונה בסוף השלב. טענה זו נכונה מכיוון שכל זוג שכנים בעלי ביטים אחרונים שונים, אחד מהם מקבל את המזהה של השני או שאחד מהם נמחק מהשלב (כן זו סוג של רמאות)

אינדוקציה: נניח כי הטענה נכונה לכל שלב $i-1$ עבור $i > 1$ ונוכיח שהיא נכונה לשלב i.
בסוף שלב $i-1$ אין קשתות בין צמתים בעלי ביט i שונה, אחרת צומת נמחק או שקלאסטר אחד מצטרף לשני.

6.3 יעילות

ישנם $O(logn)$ שלבים, $O(log^2 n)$ צעדים בכל שלב ו־ $O(logn)$ צבעים, כך שישנם $O(log^4 n)$ סיבובים.
הסיבוכיות של כל צעד היא $O(log^3 n)$ כך שבסך הכל סיבוכיות הזמן היא $O(log^7 n)$.