Network Algorithms – Lecture 4

קבוצה ב"ת מקסימלית (Maximal Independent Set – MIS)

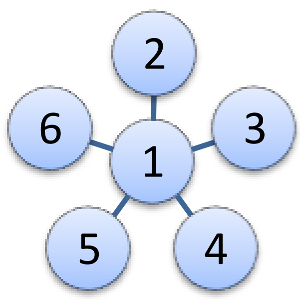
הגדרה – קבוצה ב"ת בגרף G=(V,E) הינה קבוצת צמתים $S \subseteq V$ שאין ביניהם אף קשת.

כלומר – $(SxS) \cap E = \emptyset$.

קבוצה ב"ת הינה **מקסימלית**, אם לכל צומת $v \in V \setminus S$, $S \cup \{V\}$ אינה ב"ת (לא ניתן לצרף לה עוד צומת כך שתכונת האי תלות תשמר).

קבוצה ב"ת מקסימום (maxIS)

הגדרה – קבוצה ב"ת מקסימלית שגודלה מירבי בגרף.



- בתרשים צומת 1 הינה קבוצה מקסימלית, בעוד שקבוצה ב"ת מקסימום הינה {2,3,4,5,6}

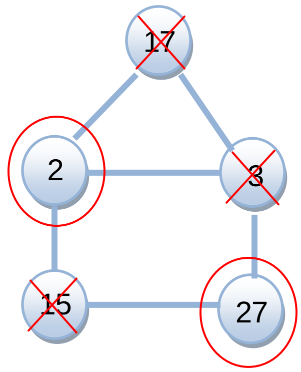
משפט – ישנו אלגוריתם פולינומי לקירוב MaxIs עד כדי מקדם $n^{1-\varepsilon}$ עבור $\varepsilon > 0$ קבוע ⇔ P=NP. כלומר קירוב הבעיה הינה בעיה NP – קשה.

לעומת זאת מציאת MIS ניתן למצוא בזמן לינארי ע"י אלגוריתם חמדן –

1. בחר צומת v, הוסף לMIS.
2. סלק את v וכל שכניו.
3. אם הגרף לא ריק חזור ל1.

אלגוריתם מבוזר ל MIS

תחילה - ישנה בעיה של א-סימטריה במציאת MIS. זאת משום שאין סיבה מוגדרת מדוע לבחור צומת מסוים לMIS לעומת צומת אחר (משכניו). לכן על מנת לטפל בבעיה זו אנו מקצים id עבור כל צומת ובוחרים צמתים עפ"י מינימום מקומי.



<u>האלגוריתם</u> -

1. כל צומת שולח לשכניו את הid שלו, אם הוא מינימום מקומי , הוא נכנס לMIS ומשדר זאת לשכניו.
2. הצומת שנכנס לMIS וכל שכניו "יוצאים מהמשחק".
3. צמתים ש"שרדו" חוזרים ל1.

<u>נכונות</u> -

הקבוצה ב"ת –

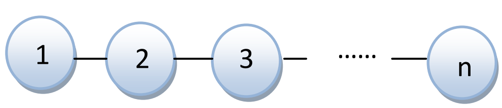
- לא ייתכן ששני שכנים יכנסו לקבוצה באותה איטרציה , כי המזהים ייחודיים , ואין יותר מ מינימום מקומי אחד.
- לא ייתכן ששני שכנים יכנסו באיטרציות שונות , כשנכנס אחד, כל שכניו יוצאים .

הקבוצה מקסימלית – כי לא עוצרים כל עוד נותר צומת בגרף.

<u>סיבוכיות זמן</u> -

<u>מקרה גרוע</u> - ייקח n/2 איטרציות

לכן במקרה הגרוע נקבל $\Omega(n)$.



<u>חסם עליון</u> – O(n) איטרציות מפני שבכל איטרציה לפחות צומת אחד (מינימום גלובלי) נכנס לMIS.

כלומר – $\theta(n)$.

<u>משפט Turan (טוראן)</u>

לכל גרף G=(V,E) עם דרגה ממוצעת $\bar{d} = \frac{2|E|}{|V|}$ עש קבוצה ב"ת בגודל של לפחות $\frac{n}{\bar{d}+1}$

<u>הוכחה</u> –

תחילה נזכר <u>באי שיוויון הממוצעים</u> -

ממוצע חשבוני – $AM = \frac{1}{n}\sum_{i=1}^{n} a_i$

ממוצע גיאומטרי – $GM = \left(\prod_{i=1}^{n} a_i\right)^{1/n}$

ממוצע הרמוני – $HM = \frac{n}{(\sum_{i=1}^{n} 1/a_i)}$

כאשר -

$$AM \geq GM \geq HM$$

כלומר – $\sum \frac{1}{a_i} \geq \frac{n^2}{\sum a_i} = \frac{n}{\bar{a}}$

<u>הוכחת משפט Turan</u> – נבחר לעל צומת v, מספר $x_v \in [0,1]$ **באקראי** בהתפלגות אחידה.

נכניס לMIS את כל צמתי המינימום המקומיים. מאחר ובוחרים מספר x_v אקראי ורציף ההסתברות ששני מספרים יהיו שונים הינה 1.

<u>נחשב תוחלת גודל הקבוצה</u> -

הסיכוי שצומת מסוים v יכנס לקבוצה הינו בדיוק $\frac{1}{d(v)+1}$

כעת, מלינאריות התוחלת, נקבל, מאי שיוויון הממוצעים , כי – . $E[\#\, nodes\ in\ s] = \sum_{v\in V} \Pr[v \in S] =$

מכיוון שהתוחלת לפחות $\frac{n}{\bar{d}+1}$ אזי קיימת קבוצה ב"ת, שגודלה לפחות - $\frac{n}{\bar{d}+1}$

אלגוריתם לMIS (בהשראת Turan).

1. כל צומת בוחר לו id באקראי ושולח לשכניו.
2. כל מינימום מקומי מודיע לשכניו (ונכנס לMIS).
3. כל צמתי המינימום המקומי ושכניהם "יוצאים מהמשחק".
4. צומת ששרד חוזר ל1.

בחירת id אקראי חייב להיות מוגדר טוב יותר , לכן נבחר id בצורה אוניפורמית מתוך $id \in [1, N^4]$ כלומר, עפ"י חסם האיחוד – $\Pr[id's\, are\, unique] = 1 - 1/N^2$ במקרה בו יהיו כמה עם אותו id נאמר שהאלגוריתם נכשל.

<u>נכונות</u> – הקבוצה ב"ת כמקודם ומקסימלית כי ממשיכים כל עוד הגרף לא ריק.

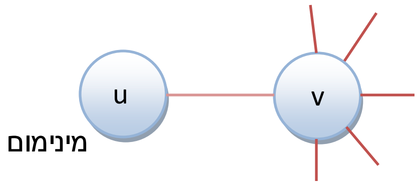
<u>ניתן לפשט את האלגוריתם אף יותר</u> –

1+2 . נוכל לומר שכל צומת v בוחר, "1" בהסתברות $\sum \frac{1}{d(v)}$ ו"0" אחרת.

3. צומת נכנס לMIS ⇔ הוא בחר "1" וכל שכניו בחרו "0".

4. צומת ששרד חוזר ל1.

כעת נספור קשתות שיורדות מהגרף ולא צמתים, נאמר שצומת u <u>מסלק</u> את צומת v שהוא שכנו אם $u = \min(N(u) \cup N(v))$, נאמר שקשת (v,w) סולקה ע"י u, אם u סילק את v (או w).



<u>למה 1</u> – ההסתברות שצומת v סולק ע"י u היא לפחות $\frac{1}{d(u)+d(v)}$.

<u>הוכחה</u> – לv יש d(v) שכנים כולל u, לu יש d(u) שכנים כולל v, לכן $d(u) + d(v) \geq |N(u) \cup N(v)|$

<u>למה 2</u> – תוחלת מספר הקשתות שסולקו הינה לפחות $\frac{|E_t|}{2}$ באיטרציה t.

<u>הוכחה</u> – $E[\#\, edge\ removed] \geq \sum_{v\in V} \sum_{u\in N(v)} \frac{d(v)}{d(u)+d(v)} = \sum_{(u,v)\in E} \frac{1}{2}\left(\frac{d(v)}{d(v)+d(u)} + \frac{d(u)}{d(v)+d(u)}\right) = \frac{|E|}{2}$

<u>למה 3</u> – בהסתברות לפחות $\frac{1}{3}$, בכל איטרציה, מסולקות לפחות $\frac{1}{4}$ מהקשתות.

<u>הוכחה</u> – נסמן $x \triangleq \Pr[at\ least\ \frac{|E|}{4}\ removed]$

כעת, עפ"י אי שיוויון מרקוב נקבל כי –

$$\frac{|E|}{2} \leq E[\#\, edges\ removed] \leq x|E| + (1-x)\frac{|E|}{4}$$

- ⇨ $\frac{1}{2} \leq \frac{1}{4} + \frac{3}{4}x$
- ⇨ $x \geq \frac{1}{3}$

כדי שכל הקשתות ירדו מספיקים $\log_{\frac{3}{4}} |E|$ סיבובים שבכל אחד יורדות לפחות $\frac{1}{4}$ מהקשתות.

<u>בתוחלת</u> – מספיקים $3\log_{\frac{n}{3}} |E|$ סיבובים.

<u>משפט</u> – האלגוריתם מסיים לבנות MIS בזמן O(logn) בהסתברות גבוהה.

לפני שנוכיח את המשפט נזכר בחסם צ'רנוף –

נניח $x_1,\dots,x_n$ משתני ברנולי (0,1) ב"ת. נסמן $X = \sum_{i=1}^{n} x_i$ ונסמן $\mu = E[X]$ אזי –

$$(\delta > 0) \qquad \Pr[X > \mu(1+\delta)] < e^{\frac{-\mu\delta^2}{4}}$$

<u>הוכחת המשפט</u> –

נסתכל על ריצה של– $3(1+\delta)\log_{\frac{3}{4}}|E|$ איטרציות.

לאיטרציה i נגדיר מ"א X_i שיהיה "1" אם מספר הקשתות שירגו באיטרציה הוא לפחות $\frac{1}{4}$ ממספר הקשתות בתחילת האיטרציה ו – "0" אחרת.

לפי למה 3, $\Pr[X_i = 1] \geq \frac{1}{3}$. ה-$X_i$-ים ב"ת כי בוחרים id מחדש בכל איטרציה.

לכן, לפי צ'רנוף, ההסתברות שיש פחות מ $\log_{\frac{3}{4}}|E|$ סיבובים עם ירידה קטנה מ $\frac{1}{4}$ היא לכל היותר –

$$1 - exp\left(-2(1+\delta)\log_{\frac{3}{4}}|E| \cdot \frac{\delta^2}{4}\right) \geq 1 - \exp\left(\frac{-ln|E|}{ln\left(\frac{4}{3}\right)} \cdot \frac{\delta^2}{2}\right) = 1 - |E|^{\frac{-\delta^2}{2\ln\left(\frac{4}{3}\right)}}$$

כלומר, בהסתברות גבוהה (δ גדול כרצוננו) לאחר, O(log|E|) = O(logn) איטרציות, הגרף נותר ללא קשתות.

צביעת צמתים

צריך לתת לכל צומת "צבע" (מספר טבעי) כך שאין צמתים שכנים בעלי אותו צבע.

המטרה למזער את מספר הצבעים בגרף

- בעיה זו הינה NP-קשה ואף קשה לקירוב עד כדי $n^{\frac{1}{4}}$.
- אם הדרגה המקסימלית של צומת הינה Δ, אזי $\Delta + 1$ צבעים תמיד מספיקים (ע"י אלגוריתם חמדן).

<u>אלגוריתם 1</u> –

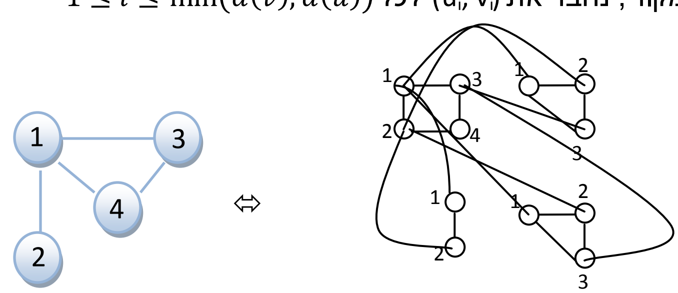
1. I=1.
2. בחר MIS.
3. צבע MIS בצבע i.
4. הסר MIS, $i \leftarrow i + 1$. לך ל1.

<u>ניתוח</u> – בכל איטרציה, דרגת כל צומת ששורד יורדת ב1. לכן אחרי $\Delta + 1$ איטרציות, סיימנו. (כאשר כל איטרציה עולה O(logn)).

<u>אלגוריתם 2</u> –

לכל צומת v, אנו ממציאים d(v)+1 צמתים וירטואליים (מחוברים בקליק).

אם הייתה קשת (u,v) בגרף המקורי, נחבר את (u_i, v_i) לכל $1 \leq i \leq \min\big(d(v), d(u)\big)$



מריצים MIS על הגרף הווירטואלי (כל צומת פיזי עושה את כל החישובים הוירטואליים על הצמתים שלו) ומספר הצמתים הווירטואליים לכל היותר n^2, לכן O(logn) סיבובים מספיקים (בהסתברות גבוהה).

נשים לב כי אם גודל ההודעה קבוע, סיבוכיות הזמן זהה לאלגוריתם 1.